The roles of climate, geography and natural selection as drivers of genetic and phenotypic differentiation in a widespread amphibian Hyla annectans (Anura: Hylidae)

Wei, Shichao

2020-10

Wei , S , Li , Z , Momigliano , P , Fu , C , Wu , H & Merilä , J 2020 , ' The roles of climate, geography and natural selection as drivers of genetic and phenotypic differentiation in a widespread amphibian Hyla annectans (Anura: Hylidae) ' , Molecular Ecology , vol. 29 , no. 19 , pp. 3667-3683 . https://doi.org/10.1111/mec.15584

http://hdl.handle.net/10138/333047
https://doi.org/10.1111/mec.15584

acceptedVersion

The roles of climate, geography and natural selection as drivers of genetic and phenotypic differentiation in a widespread amphibian *Hyla annectans* (Anura: Hylidae)

Shichao Wei[1], Zitong Li[2], Paolo Momigliano[2], Chao Fu[1], Hua Wu[1]*, & Juha Merilä[2]

[1]Institute of Evolution and Ecology, School of Life Sciences, Central China Normal University, 152 Luoyulu, Hongshan District, 430079 Wuhan, China

[2]Ecological Genetics Research Unit, Organismal and Evolutionary Biology Research Programme, Faculty of Biological and Environmental Sciences, University of Helsinki, P.O. Box 65, FI-00014 Helsinki, Finland

Corresponding author: Hua Wu (wuhua@mail.ccnu.edu.cn, Telephone: 86-027-67867021; fax:86-027-67861147); Institute of Evolution and Ecology, School of Life Sciences, Central China Normal University, 152 Luoyulu, Hongshan District, 430079 Wuhan, China

Running title: Phylogeography of *Hyla annectans*

Abstract

The role of geological events and Pleistocene climatic fluctuations as drivers of current patterns of genetic variation in extant species has been a topic of continued interest among evolutionary biologists. Nevertheless, comprehensive studies of widely distributed species are still rare, especially from Asia. Using geographically extensive sampling of many individuals and a large number of nuclear single nucleotide polymorphisms (SNPs), we studied the phylogeography and historical demography of *Hyla annectans* populations in southern China. Thirty-five sampled populations were grouped into seven clearly defined genetic clusters that closely match phenotype-based subspecies classification. These lineages diverged 2.32–5.23 million years ago, a timing that closely aligns with the rapid and drastic uplifting of the Qinghai-Tibet Plateau and adjacent southwest China. Demographic analyses and species distribution models indicate that different populations of this species have responded differently to past climatic changes. In the Hengduan Mountains, most populations experienced a bottleneck, whereas the populations located outside of the Hengduan Mountains have gradually declined in size since the end of the last glaciation. In addition, the levels of phenotypic and genetic divergence were strongly correlated across major clades. These results highlight the combined effects of geological events and past climatic fluctuations, as well as natural selection, as drivers of contemporary patterns of genetic and phenotypic variation in a widely distributed anuran in Asia.

Keywords: phylogeography, climatic fluctuations, population divergence, SNP, natural selection

Introduction

Contemporary patterns of genetic variation across different geographic areas are affected by historical factors (Avise 1994; Hewitt 2004). Geological events such as the formation of mountain ranges and river systems can generate physical barriers to dispersal, fragmenting once connected habitats, hence resulting in allopatric divergence and speciation (Che *et al.* 2010; Chaves *et al.* 2011). Past climatic fluctuations, particularly those during the late Pleistocene, were important drivers of current distributions, genetic diversification, and demographic fluctuations of many temperate species and communities (Hewitt 2000; Hewitt 2004). During glacial periods, many taxa retreated into refugia and subsequently underwent range expansions during postglacial periods in response to the availability of newly formed habitats (Hewitt 2004). Hence, contemporary patterns of genetic variation within temperate zone species have been likely influenced by both paleogeological events and Pleistocene climatic fluctuations (Hewitt 1996; Kumar & Kumar 2018; Li *et al.* 2018).

Genetic methods have been widely used to investigate the effects of climate and geography in driving current patterns of genetic and phenotypic variation within species (Avise 2000; Hewitt 2000; Leinonen *et al.* 2013; Kumar & Kumar 2018). Phylogeographic studies were initially based on mitochondrial DNA (mtDNA; Avise *et al.* 1987), and subsequently complemented with nuclear markers (e.g. Yan *et al.* 2013; Li *et al.* 2018). Given the now well-known limitations of evolutionary inferences based on mtDNA (Avise 2000; Ballard & Whitlock 2004; Guo *et al.* 2019), next generation sequencing (NGS) methods have largely replaced mtDNA and microsatellite markers in phylogeographic and population genetic investigations of non-model organisms (Avise 2009; McCormack *et al.* 2013). Although phylogeographic studies using NGS data have become increasingly common, most have focused on European and North American areas (e.g. Newman & Austin 2016; Dufresnes *et al.* 2019), while large-scale phylogeographic studies based on NGS data from Asia are still relatively rare (but see: Zhao *et al.* 2013; Zhou *et al.* 2016; Puckett *et al.* 2016; Jiang *et al.* 2018; Wang *et al.* 2018; Feng *et al.* 2019)

Southern China provides an interesting area for phylogeographic studies due to its unique geophysical conditions and abundant biodiversity (Yan *et al.* 2013; Li *et al.* 2015b; Li *et al*. 2018). The geological features of Chinese mainland have been remodeled by the uplift of the Himalayas and the Qinghai-Tibet Plateau (QTP; Harrison *et al.* 1992; Zhang 1999). Presently, these areas are characterized by many high-elevation mountains, plateaus and river systems such as the Hengduan Mountains, Yunnan-Guizhou Plateau, and Yangtze River (Zhang 1999). These geographic barriers have likely played important roles in driving the genetic and phenotypic divergence of the species native to this region (Che *et al.* 2010; Yan *et al.* 2013). Over the East Asian continent temperatures during the LGM were 2–4 ℃ colder than today (Weaver *et al.* 1998; Ju *et al.* 2007), but unlike in Europe and North America, most areas in southern China were not covered by ice sheets during the Pleistocene (Shi *et al.* 1986; Liu 1988) – with the exception of the Hengduan Mountains (Zheng *et al*. 2002). Thus, Quaternary climatic fluctuations might have had less impact on patterns of genetic variation in southern China compared to Europe and North America, and their impact within regions of Southern China might have been heterogeneous (Wang & Ge 2006; Gao *et al.* 2011; Yan *et al.* 2013).

Amphibians have been identified as good models for studying the factors that shape the patterns of genetic variation and differentiation, mainly for two reasons. First, because of their limited dispersal ability, they display very high levels of population genetic structuring compared to other animal classes (Ward *et al*. 1992; Zeisset & Beebee 2008; Sánchez-Montes *et al.* 2018). Second, as ectotherms they are sensitive to climatic conditions, and are thereby considered to be good indicators of climate change, both past and present (Bossuyt & Milinkovitch 2001; Graham *et al.* 2004; Kozak & Wiens 2010).

The Jerdon's tree frog *Hyla annectans* (Anura: Hylidae) is widely distributed in Asia south of the Himalayas. In southern China, it occurs in low-to-medium elevation (ca. 580–2,500 m above sea level) forests (Fei *et al.* 2009). Currently, five subspecies (*viz. H. a. gongshanensis*, *H. a. tengchongensis*, *H. a. jingdongensis*, *H. a. chuanxiensis*, and *H. a. wulingensis*) with disjunct geographic distributions are recognized (Fei *et al*. 2009). The subspecies display phenotypic divergence in number and shape of

black spots on their flanks (Fei *et al.* 2009): such divergence indicates that these traits may have been subject to divergent sexual and/or natural selection. Given that the most recent common ancestor (MRCA) of *H. annectans* dates back to the mid Pliocene (~4 Mya, 95% CI: 3–5 Mya; Li *et al.* 2015a), the species has been exposed to the intense uplift of the Qinghai-Tibetan Plateau (QTP) and adjacent southwest China (Cui *et al*. 1996; Sun *et al*. 2011) and subsequent climatic oscillations. Hence, it is an ideal amphibian model system to study the effects of past geomorphological events and historical climatic fluctuations on phylogeography and historical demography.

The primary aim of this study was to investigate the impacts of the past geomorphological events and Pleistocene climatic fluctuations on the patterns of genomic differentiation and historical demography in *H. annectans*. We screened thousands of genome wide genetic markers in a large number of samples covering most of the species distribution area in China, and subjected the data to various population genomic analyses, species distribution modelling, as well as analyses of historical demography. In addition, we tested for effects of natural selection on phenotypic traits, and whether the levels of (presumably neutral) genetic divergence among populations predicts levels of phenotypic divergence. Hence, the results were expected to yield insights as to how past geological events, climatic fluctuations and natural selection have shaped the distribution of genetic and phenotypic variation in an amphibian distributed over a large geographic area.

Materials and methods

Sampling and DNA extractions

The sample collection of *H. annectans* was planned based on the maps provided in Fei *et al.* (2009). In total, we obtained 349 samples from 35 sites collected throughout the species' distribution range in China (Figure 1, Table S1). Ten adult specimens per site were collected, with the exception of location "20", from where nine specimens were obtained (Table S1). Muscle tissue was taken from each specimen and preserved in 99% ethanol in the field, and later transferred to a -20 ℃ freezer in the

Molecular and Behavioral Ecology Research Group Laboratory, Central China Normal University, Wuhan. Genomic DNA was extracted using a standard CTAB protocol (Hanania *et al.* 2004). DNA concentration and quality were assessed using a ND-1000 spectrophotometer (NanoDrop, Wilmington, DE, USA); DNA quality was also checked on 1% agarose gels with lambda DNA standard.

High-throughput sequencing and single nucleotide polymorphism data assembly

We used the high-resolution Specific Length Amplified Fragment Sequencing (SLAF-seq) strategy for large-scale *de novo* SNP discovery and genotyping (Sun *et al.* 2013). The genome of *Xenopus tropicalis* (GenBank assemble accession: GCA_000004195.2) was chosen as a reference for running an *in silico* digestion to determine an appropriate combination of restriction enzymes. Appropriate enzymes should result in a large number (> 100,000) of unique SLAF markers that are randomly distributed throughout the training genome and contain a low proportion of repeat sequences. Based on the training results, we chose a combination of *HaeIII* and *Hpy166II* (New England Biolabs, NEB) restriction enzymes with a size-selection window of 414–444 bp, which was expected to yield approximately 110,000 SLAF tags in *X. tropicalis*. These enzymes were used to digest the genomic DNA of *H. annectans* for SLAF-seq library construction. Genomic DNA of each sample was digested with *HaeIII* and *Hpy166II* (New England Biolabs, NEB), and a dATP was used to add a single nucleotide (A) with Klenow Fragment (3'→5' exo-) (NEB). Duplex Tag-labeled Sequencing adapters (PAGE purified, Life Technologies, USA) were ligated using T4 Ligase to the A-tailed DNA. The PCR reactions were run using diluted restriction-ligation samples, dNTPs, Q5® High-Fidelity DNA Polymerase, and forward (5'-AATGATACGGCGACCACCGA-3') and reverse (5'-CAAGCAGAAGACGGCATACG-3'; PAGE purified, Life Technologies, Beijing) primers. PCR products were purified using Agencourt AMPure XP beads (Beckman Coulter, High Wycombe, UK) and pooled. The pooled samples were checked by electrophoresis in a 2% agarose gel, and fragments varying in length from 414 to 444 bp (with indices and adaptors) were isolated using a Blue Pippin (Sage Science, Beverly, MA). The purified products

were submitted for paired-end 100-bp sequencing on the Illumina HiSeq 2500 system (Illumina, Inc; San Diego, CA, USA) according to the manufacturer's guidelines.

Data processing and SNP calling

Adaptor contamination, primer contamination, and low quality reads were present in the raw sequence reads. FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to run an initial quality check on the raw data, and low quality reads (N content > 10%, more than 50% of bases with quality values < 10) were removed. Given the paucity of genomic resource for *H. annectans* and related species (e.g. Hylidae), we clustered all the paired-end reads into SLAF loci with clear index information based on sequence similarity above 90% using BLAT (Kent 2002) and concatenated all loci into a "fake" reference genome. For each locus the reference sequence was selected based on maximum sequencing depth of the corresponding SLAF tag. We used these matched sequences as our reference for sequence alignment and SNP calling. High-quality reads were mapped onto this reference using BWA-MEM (Li & Durbin 2009). The mapped reads were then sorted and duplicate reads were removed using PICARD-TOOLS v.1.67 (http://broadinstitute.github.io/picard/). Local realignment around the indel-regions was performed using RealignerTargetCreator and IndelRealigner in GATK (Genome Analysis Toolkit; McKenna *et al.* 2010). Since different variant calling pipelines may be prone to unique biases and provide inconsistent results (O'Rawe *et al.* 2013; Clevenger *et al.* 2015), we called variants using both the mpileup command in SAMTOOLS v.1.1 (Li *et al.* 2009) and GATK UnifiedGenotyper with default settings. We selected the concordant common sites identified by both GATK and SAMTOOLS using the SelectVariants package with default settings in GATK. Variant filtering was performed following the 'Best Practices' workflow developed by the GATK team (McKenna *et al.* 2010). Sequencing depths of each sample were calculated using the 'Depth of Coverage' module of GATK after removing indels with the SelectVariants package in GATK. The number of SLAF tags varied from 90,581 to 162,334 across individual samples, and a total of 1,075,515 SLAF tags and 2,303,646 biallelic SNPs were retained. For phylogenetic and population genetic analyses, we excluded SNPs with allele count < 35

and with missing data over 20% across all individuals. Only one SNP per locus was retained. Individuals with more than 40% missing data were removed (Zhao *et al*. 2016). The final filtered dataset included 8,420 informative SNPs. For divergence time estimation and TREEMIX analyses, only SNPs with MAF > 0.05 and with less than 5% or 10% missing data, respectively, were retained. For demographic analyses (i.e. STAIRWAY PLOTS) no missing data were allowed, and the data were not filtered for MAF to avoid distorting the allele frequency spectra. More details on the different datasets are given in Supplementary Table S2.

Phylogenetic inference

A phylogeny of *H. annectans* populations was first estimated by constructing a Neighbor Joining (NJ) tree based on maximum composite likelihood with 10,000 bootstrap replicates using the MEGA X software (Kumar *et al.* 2018), with *H. sanchiangensis* as an outgroup. We estimated divergence times among lineages under the Multispecies Coalescent using the SNAPP v1.4.1 (Bryant *et al*. 2012) plugin of BEAST v2.4.4 (Bouckaert *et al*. 2014) with a molecular clock model (Stange *et al.* 2018). Since SNAPP is too computationally demanding to analyze all our individuals, we used a smaller dataset of 72 individuals generated by randomly sampling two individuals from each site and from the outgroup. This dataset included a total of 2,183 SNPs with < 5% missing data. We used the time to most recent common ancestor (tMRCA) of *H. annectans* (set as a lognormal distribution with 4 Mya ± 0.14) and tMRCA between *H. annectans* and the outgroup *H. sanchiangensis* (set as a lognormal distribution with 11.6 Mya ± 0.18) as calibration nodes. We obtained these priors from a time-calibrated phylogeny of the genus *Hyla* based on mitochondrial and nuclear genetic data with three fossil calibration points (Li *et al.* 2015a). In SNAPP, we ran three independent analyses with 1,000,000 MCMC iterations. We thinned each chain by sampling every 1,000 trees to reduce serial correlation and checked the convergence of the MCMC and effective sample sizes (above 200) in TRACER v.1.7 (Rambaut *et al*. 2018). We combined the results from the three independent chains in LOGCOMBINER v2.4.4 (Bouckaert *et al.* 2014). We used the program DENSITREE v.2.2.6 (Bouckaert *et al.* 2014) to visualize

the SNAPP trees after discarding the first 10% of each MCMC chain as burn-in. Finally, we summarized the maximum-credibility trees with median heights in TREEANNOTATOR v.2.4.4 (Drummond & Rambaut 2007).

Molecular diversity and genetic structure

We examined the patterns of genetic structuring among *H. annectans* populations with two different methods. Firstly, we used the fast variational Bayesian algorithm implemented in the software fastSTRUCTURE (Raj *et al.* 2014). Values for K = 2–15 were tested to determine the optimal number of clusters (K) using the Bayesian model selection criterion provided by fastSTRUCTURE. We ran the analyses for the best-supported number of clusters by *a)* using the *chooseK.py* program (Raj *et al.* 2014), which infers the best fitting model as the number of K that maximizes the marginal likelihood low bound (LLBO), and *b)* running a fivefold cross-validation and choosing the value of K that minimized prediction error. To visualize population structure, we used the web application POPHELPER (Francis 2017). Secondly, we ran a principal component analysis (PCA) based on the sample covariance matrix of the SNP data (Patterson *et al.* 2006) using the R package ADEGENET (Jombart 2008).

After defining the genetic lineages of *H. annectans* on the basis of genetic clustering and phylogenetic analyses, we calculated genetic diversity indices including the expected (*He*) and observed heterozygosity (*Ho*) for each lineage using the R package ADEGENET (Jombart 2008). We estimated pairwise F_{ST} among genetic clusters in ARLEQUIN v.3.5.2.2 (Excoffier & Lischer 2010); 10,000 permutations were run to test for statistical significance.

Importance of environmental and geographical factors in explaining genetic differentiation

We plotted Slatkin's linearized F_{ST} (Slatkin 1995) against geographic distance to determine whether the observed patterns of genetic differentiation conform to Isolation by Distance model (IBD), and tested for a correlation between genetic and geographic distance matrices using a Mantel test with the ADE4

package in R. We estimated geographic distances calculated in ARCMAP implemented in ARCGIS Desktop version 10.3 (ESRI) based on latitude and longitude data for sampling site, and extracted the values using the "point distance" function.

In addition, we used a distance-based redundancy analysis (dbRDA) to test the effects of environmental and geographical factors on explaining genetic differentiation of *H. annectans* populations. We set the pairwise F_{ST} as response variable. To obtain geographic explanatory variables, we computed a Euclidian distance matrix from the Cartesian coordinates for each sampling site using the "dist" function and performed the "pcnm" function (permutations = 1000) on this matrix to obtain uncorrelated vectors. We then selected the positive eigenvectors as spatial variables as they were positively correlated with the geographic distance. The first three positive vectors (GEO1, GEO2, GEO3) were retained and used to run the dbRDA analysis. As environmental explanatory variables, we used four climatic variables (BIO1, BIO2, BIO12 and BIO14) that minimized collinearity. We estimated the relative contributions of both geographic and environmental variables and their intersection by variance partitioning. We additionally applied the dbRDA to detect IBD, considering the widespread concerns about the reliability of Mantel tests (Kierepka & Latch 2014). All those analyses were performed by the "capscale" and "anova.cca" functions in R package VEGAN (Oksanen *et al*. 2019).

Demographic analyses

We estimated past changes in effective population size (N_e) for each sampling location with STAIRWAY PLOTS derived from folded SFS data (Liu & Fu 2015) in order to evaluate the effects of paleoclimatic changes. Since F_{ST} values were significant among most of the population pairs, we estimated the SFS for every single sampling location from a subset of biallelic SNPs with no missing data using ANGSD (Korneliussen *et al*. 2014), which resulted in 35 SNP datasets (Table S2). To construct STAIRWAY PLOTS, we used the default 2/3 of the data for training and [(nseq-2)/4, (nseq-2)/2, 3*(nseq-2)/4 and (nseq-2)] as the number of random breakpoints (nrand, where nseq indicates the

number of sequences). We set generation time and mutation rate to two years (Liao & Lu 2010) and 1.552×10^{-9} substitutions per site per generation (Sun *et al.* 2015), respectively.

We reconstructed migration events among *H. annectans* populations using TREEMIX v1.12 (Pickrell & Pritchard 2012) based on 2,118 informative SNPs. The model scenario was specified as follows: we set the number of migration events to be from 1 to 20 (m=1–20), block size to 50, and *H. sanchiangensis* as the outgroup for the purpose of rooting. To evaluate the optimal number of migrations, we calculated the variance of relatedness between populations explained by the model using *TreemixVarianceExplained.R* (https://github.com), with over 99.8% of variance suggesting a reliable model (Pickrell & Pritchard 2012).

Species distribution modelling (SDM)

We generated SDMs for four time periods: the present time, the Mid-Holocene (5–7.5 kya), the Last Glacial Maximum (LGM; about 21 kya) and the Last Interglacial (LIG; about 120–140 kya) to investigate the possible influence of climatic changes on the distribution of *H. annectans*, using MAXENT v.3.3.3e (Phillips *et al.* 2006). We obtained the occurrence data of *H. annectans* used to build the SDMs in this study from four sources: samples used in this study, two literature records (Liao & Lu 2010; Shen 1996), the Global Biodiversity Information Network GBIF (http://www.gbif.org), and VertNet (http://vertnet.org/; Table S3). Firstly, we removed data with no detailed locality information, imprecise GPS coordinates, obviously erroneous location (i.e. located in water bodies), as well as duplicate data. We only retained data at a resolution higher than 5 km, which corresponds to the resolution of the climatic data of each grid cell with size of 2.5 arc minutes (approximately 5 km). This resulted in a total 119 occurrence data points (Table S3). They were further assigned to main lineages (*viz*. lineage E [n=23], C [n=26], N1 [n=15], N2 [n =12], and W [n=43]) as inferred by the phylogeographic analyses (see Results). We searched for the best combinations of feature classes (determining the shape of the response curves) and regularization multipliers (determining the penalty

for adding parameters in the model) by evaluating model scores based on the Akaike information criterion (AICc). We used the ENMeval package (Muscarella *et al*. 2014) to identify the best model with the "ENMevaluate" function in R. The model feature types used were 'L', 'H', 'LQ', 'LQH', 'LQHP', 'LQHTP' (where: L = linear, Q = quadratic, H = hinge, P = product and T = threshold) and regularization (RM) values (0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4). For a proper evaluation, we calibrated SDMs using a random subset of 75% of the sampling sites; the remaining 25% were reserved to test the validity of the models. We used 19 bioclimatic layers as environmental predictors at 30-arcsec (~1 km) resolution, which we downloaded from the WorldClim database (http://www.worldclim.org/; Hijmans *et al*. 2005). To avoid multicollinearity, we selected BIO (bioclimatic variables) parameters using PCA. The results showed that the variance of the climate in the study area could be explained by four principal components (PCs) that captured 90% of the variance in the data. Thus, we selected one representative BIO parameter per PC (BIO1 = Annual Mean Temperature, BIO2 = Mean Diurnal Range, BIO12 = Annual Precipitation and BIO14 = Precipitation of Driest Month, Table S4) to create the SDMs. The three general circulation models (GCMs) used to generate Mid-Holocene and LGM climate scenarios were the CCSM4, MIROC-ESM (Watanabe *et al*. 2011) and MPI-ESM-P models available from the WorldClim database (http://www.worldclim.org/). Only one GCM of the LIG period was available. We used ARCGIS v.10.3 to manipulate and visualize the spatial environmental data and model output.

We employed the Mobility-oriented parity (MOP; Owens *et al*. 2013) analysis to assess if the study areas had similar environmental conditions currently, during the LGM and during the LIG, and if extrapolation risks exist. We used the 10% as a subsampling percentage for study area in the current climate. We performed the analysis using the KUENM package (Cobos *et al*. 2019) in R.

Niche divergence

We compared SDMs products for the five main lineages (E, C, N1, N2 and W) separately, to evaluate niche divergence in their predicted niche distribution by using ENMTools (Warren *et al*. 2008). We utilized two metrics for calculating niche divergence from the MAXENT: Schoener's *D* (Schoener

1968) and Warren's *I* statistic (Warren *et al*. 2008). Both metrics range from 0 to 1, with 0 corresponding to identical niches and 1 representing no niche divergence between the two compared groups.

Morphological analyses

We also tested for differences between the main identified lineages and/or recognized subspecies (Fei *et al*. 2009). For each collected individual, we counted the number of distinct round black spots on the right (posterior) side of the body, as this is a taxonomically diagnostic character used to demarcate different subspecies (Fei *et al*. 2009). We measured the snout-vent length (SVL) to the nearest 0.01 mm with digital calipers and weighed the specimens to the nearest 0.1g with electronic digital balance. In total, we grouped 339 individuals (10 individuals from population "2" were not measured) according to their genetic cluster and compared the mean values of black spot numbers across clusters using a Kruskal-Wallis test, as trait values were not normally distributed. We also compared mean size and weight of individuals using a parametric ANOVA, as these traits were normally distributed. We performed all statistical using the SPSS software (SPSS 22.0, SPSS Inc, Chicago, IL, USA) and tested for significance at an alpha level of 0.05. We visualized the relationship between the number of black spots and the seven genetic clusters in R v.3.2.2 (R Core Team 2014).

Q_{ST}-F_{ST} comparison

We conducted Q_{ST}-F_{ST} comparisons to explore whether the degree of phenotypic differentiation in three traits (number of spots, snout-vent length, and weight) exceeded neutral expectation – which would indicate differentiation driven by natural selection – by using the R packages RAFM and DRIFTSEL (Ovaskainen *et al*. 2011; Karhunen *et al*. 2013). A Q_{ST} is a metric equivalent to F_{ST} – while the latter is estimated from genetic marker data and reflects the degree of neutral genetic differentiation, the former is derived from phenotypic data reflects the degree of differentiation in quantitative traits (e.g. Leinonen *et al*. 2013). A Q_{ST} significantly larger than F_{ST} would be indicative of differentiation in given quantitative trait exceeding neutral expectation, and hence, that the divergence in trait values is driven

by natural selection (Leinonen *et al*. 2013). Compared to conventional frequentist approaches, the RAFM/DRIFTSEL uses MCMC-based Bayesian algorithms to account for patterns of relatedness among populations, as well as ancestral genetic correlations among the traits of interest, hence the power to detect signatures of selection from data with small sample sizes is stronger than the conventional Q_{ST}-F_{ST} comparisons (Ovaskainen *et al*. 2011). First, the RAFM software calculated the F_{ST} and the genomic relatedness among the 35 populations based on the 8,420 SNPs. Next, we used DRIFTSEL to estimate the Q_{ST} of the three traits: number of black spots, snout-vent length and weight, and also to perform the comparison between Q_{ST} and F_{ST}. The final output of the DRIFTSEL analysis is a so-called S-statistic. S-values close to zero are indicative of stabilizing selection; those close to one indicate directional selection; and values close to 0.5 are consistent with evolution due to drift. Following the testing criteria proposed in Karhunen *et al*. (2014), $S > 0.95$ implies that a quantitative trait has evolved under divergent selection at the 95% credibility level, whereas $S < 0.05$ would imply stabilizing selection at the same credibility level. The default non-informative priors were used in the DRIFTSEL analyses. 15,000 Markov Chain Monte Carlo (MCMC) samples of the posterior distribution were simulated, and the first 5,000 were discarded as burn-ins. The remaining were stored in every 10th iteration, so that eventually 1,000 MCMC samples were used for calculating S-statistics.

We estimated the added variance component for the three abovementioned phenotypic traits to see whether the degree of phenotypic divergence is predictable from the degree of genetic divergence (F_{ST}), using a standard ANOVA approach (Sokal and Rohlf 1981). This quantity named as P_{ST} (Leinonen *et al*. 2013) is similar to Q_{ST} (= the degree of genetic differentiation in quantitative traits; Leinonen *et al*. 2013) under certain assumptions (within- and among-population components of variance are not confounded by environmental effects; Brommer 2011; Leinonen *et al*. 2013, see also Discussion). Since there is no reason to assume that these quantities would be normally distributed, we used Spearman rank correlation coefficient for testing this association.

Results

Sequencing and SNP calling

We sequenced 349 individuals of *H. annectans* using an Illumina HiSeqTM2500, generating a total of 479 million paired-end reads. There were 82.78% bases with quality scores of at least 30 (Q30) and the guanine-cytosine content was 42.06%. We obtained 1,075,515 tags (or SLAFs) in total, and their average sequencing depth was 5.53 (Table S5). A total of 2,303,646 bi-allelic SNPs were obtained. After filtering, the four datasets contained 8,420 SNPs for phylogeny and structure analysis, 2,118 for estimating gene flow, 2,183 for divergence time estimation, and 3,002–14,842 SNPs (depending on the population) for analyses of historical demography.

Phylogenetic inference

Phylogenetic analyses based on 8,420 SNPs revealed seven major genetic clusters in the NJ-tree, concordant with geography (Figure 1). The eastern cluster (E) included 11 populations from the Wuling Mountains (Figure 1). The central cluster (C), which contained two sub-clusters (C1 and C2), included seven populations distributed across central and eastern Yunnan-Guizhou Plateau (Figure 1). The northern cluster (N) had two sub-clusters (N1 and N2), representing populations located in the Hengduan Mountains on the margin of the Sichuan Basin (Figure 1). The western cluster also contained two sub-clusters (W1 and W2) distributed along the Hengduan Mountains and western Yunnan-Guizhou Plateau, respectively (Figure 1).

We estimated that the lineages of *H. annectans* initiated their divergence during the Pliocene. The northern clade (N1 and N2) diverged from the other clades ca. 5.23 million years ago (Mya, with 95% highest posterior density interval (HDPI) of 4.38–6.55 Mya, Figure 2 and Figure S1). The eastern clade (E) diverged from the western and central clades (W and C) approximately 4.88 Mya (95% HDPI: 4.09–6.15 Mya). Lineages N1 and N2 diverged approximately 4.44 Mya (95% HDPI: 3.60–5.80 Mya) and subsequently, the western and central clades split into two clades at about 3.77 Mya (95% HDPI: 3.22–4.84 Mya). The genetic divergences within western and central clades were estimated to have occurred

at 2.32 Mya (95% HDPI: 1.63–2.90 Mya) for C1 and C2, and at 2.39 Mya (95% HDPI: 1.94–3.10 Mya) for W1 and W2 (Figure 2 and Figure S1).

Molecular diversity and genetic structure

The variation in observed and expected heterozygosity among clusters was considerable, ranging from 0.013 (population 22) to 0.134 (population 11) for observed heterozygosity, and from 0.014 (population 22) to 0.177 (population 11; Table S6) for expected heterozygosity. The highest observed and expected heterozygosities occurred in clusters E, C1 and C2, which were located in the Wuling Mountains and the East Yunnan-Guizhou Plateau (Table S6). Pairwise F_{ST} values were all statistically significant, with an average pairwise F_{ST} = 0.669, ranging from 0.016 to 0.957 ($P < 0.001$; Table S6), suggesting that *H. annectans* populations are geographically highly structured.

The Bayesian clustering algorithm implemented in fastSTRUCTURE detected a clear geographical pattern of subdivision (optimal K = 7, Figures 3a and S2). At K = 2, the eastern (E) cluster was distinct from other clusters (Figure 3a); at K = 3, the western clusters (W1 and W2) were distinct from the central (C1 and C2) and northern (N1 and N2) clusters; at K = 4, the central clusters separated from the northern clusters; at K = 5, the northern clusters were divided into N1 and N2; at K = 6, the central cluster was divided into two sub-clusters C1 and C2, where population 15 was an admixture between the two sub-clusters; at K = 7, the W2 cluster showed signs of admixture with an unsampled population. In addition, within the western cluster, cluster W2 was indicated to be an admixture between western and central clusters at K = 3-6; three populations (‘9’, ‘10’ and ‘11’) from the eastern cluster showed signs of admixture with the central cluster (C1) at K = 2-7. The clustering analyses using the simple models showed an optimal choice of K = 7, which had the lowest value of cross-validation (CV) error and the highest marginal likelihood (Figure S2). While the W1 and W2 clusters were adjacent, they were clearly separated in the PCA. Similarly, while the geographic distance separating clusters N1 and N2 was very short (e.g. the distance between sampling locations ‘20’ and ‘22’ is 38 km), all analyses indicated that

they were very highly differentiated (F_{ST} = 0.956) independent genetic clusters. In summary, results from the fastSTRUCTURE, PCA, and phylogenetic analyses suggest that the 35 *H. annectans* populations sampled from southwestern China are divided into seven geographically and genetically distinct lineages (N1, N2, W1, W2, C1, C2 and E).

Importance of environmental and geographical factors in explaining genetic differentiation

The IBD test revealed a weak but significant correlation (dbRDA r^2_{adj} = 0.171, P = 0.001; Mantel test statistic r^2 = 0.015, P = 0.003; Figure S3) when all populations were included. However, a few populations deviating from the general pattern (populations in clusters N1, N2 and W1 located at opposite sides of the Hengduan Mountains) showed a very high degree of differentiation over short geographic distances, suggesting very limited gene flow despite close geographic proximity. When we excluded these deviating populations, which comprised 14 out of the 35 sampling locations, a much stronger IBD was apparent (dbRDA r^2_{adj} = 0.454, P = 0.001; Mantel test statistic r^2 = 0.343, P < 0.001; Figure S3).

The redundancy analyses revealed that the contribution of environmental variables to genetic divergence was somewhat higher than that of geographic variables (41 and 37%, respectively; Table 1). The variance partitioning test showed that the environmental and geographic variables explained 15.1% and 14.7% of the variance, respectively, whereas their intersection explained 7% of variance (Figure S4).

Demographic analyses

The STAIRWAY PLOTS revealed diverse demographic histories for populations from different lineages. Populations in lineages W1, W2, E, C1 and C2 maintained stable population sizes during the LIG (Figure 4). Most of the populations within the Eastern and Central clusters experienced a moderate population size contraction during or shortly after the LGM or the Holocene Optimum. Populations

within clusters N1, N2 and W1 showed clear signs of strong population contractions followed by expansions during the Holocene Optimum and the LGM, respectively (Figure 4a,b,c).

Although populations in the seven genetic lineages diverged and experienced different demographic histories, TREEMIX analyses (99.8% of variance explained) identified 12 migration events between populations (Figure 4h). Specifically, obvious migration events from populations in lineage C1 to the populations in lineage E, and between populations in C groups with populations in lineage W2 (Figure 4h). We also found two weakly supported migration events from populations in lineage W2 to populations in lineage N2, and the ancestor of the northern populations could be attributed to an admixture event with the W2 lineages. Interestingly, there was a clear migration event from lineage E to lineage N1 despite the long geographic distance separating them, whereas no distinct migration event was detected between populations in lineages N1 and N2, which are closest together (Figure 4h).

Species distribution modelling

The lowest AICc was assigned to the LQHP 1 model (Figure S5 and Table S7), and this was chosen to generate the projections of species distribution model. The distribution model accurately predicted the distribution area under the curve (AUC) values (mean ± SE: 0.951 ± 0.123, Table S8), indicating a good performance of the predictive models. The predicted current species distribution area was generally similar to the actual known distribution area in China (Figure 1, 6 and Figure S6). The overall suitable distribution areas of all lineages shrank gradually from the LIG to LGM (Figure 5 and Figure S6). The suitable distribution areas of the N1 lineage was indicated to have reduced significantly, especially for the northwest corner of the Sichuan Basin from the LGM to the present (Figure 5c and Figure S6). The suitable distribution area of the N2 lineage was indicated to have shrank in the south region from the LGM to the present (Figure 5d and Figure S6). In contrast, several other regions including the southern Hengduan Mountains (sites for lineage W; Figure 5e and Figure S6), Yunnan-Guizhou Plateau (sites for

lineage C; Figure 5b and Figure S6) and Wuling Mountains (sites for lineage E; Figure 5a and Figure S6) have reduced slightly in size from the LGM to the present.

The MOP analyses demonstrated that past scenarios mostly possess climates analogous to those in the current scenario in the distribution areas of *H. annectans* (Figure S7). Most of the strict extrapolation risk was present in the northern and eastern parts of the distribution maps for each period (HM, LGM, and LIG; Figure S7).

Niche divergence

Based on both Schoener's *D* and Warren's *I*, there was stronger niche divergence between the comparisons involving C and N1 lineages, E and N1 lineages, E and N2 lineages, E and W lineages than in the comparisons of between the N1 and W lineages, W and C lineages, and E and C lineages (Table 2). It is notable that the comparison of the geographically closely located N1 and N2 lineages, a moderate degree niche divergence was observed (Schoener's $D = 0.337$; Warren's $I = 0.634$, Table 2).

Morphological variation

Comparison of mean number of black spots among the seven genetic clusters revealed significant differences between the clusters (Table 3). Comparisons of mean snout-vent length and weight among the genetic clusters also indicated significant differences (ANOVAs, snout-vent length: $F_{6,319} = 3.13$, $P < 0.01$; Weight: $F_{6,319} = 9.96$, $P < 0.01$). However, only few pairwise comparisons revealed significant differences in SVL (4.67% of comparisons) and weight (33.3% of comparisons; Table S9).

Q_{ST}-F_{ST} comparison

DRIFTSEL yielded S = 0.95 for number of black spots, 0.44 for snout-vent length, 0.45 for weight and 0.84 for all three traits tested together. Hence, the only evidence for natural selection was differentiation in the number of black spots, but not for the other traits whether considered separately or together.

Likewise, the correlations between pairwise F_{ST} and pairwise P_{ST} for the three different phenotypic traits were significant for number of black spots (r_{337} = -0.53, $P = 10^{-16}$), but not for snout-vent length (r_{337} = -0.18) or weight (r_{337} = -0.09).

Discussion

We discovered that *H. annectans* in southern China consists of seven phylogenetic lineages with parapatric distributions. These lineages show high levels of genetic differentiation and also clear phenotypic divergence likely attributable to natural selection. These distinct genetic lineages diverged from each other 2.32 to 5.23 Mya, a timing that is broadly consistent with the rapid and intense uplift of the QTP and adjacent southwest China (Cui *et al*. 1996; Sun *et al*. 2011). The results further indicate that both geographic and environmental factors have contributed to the observed genetic differentiation. Demographic analyses and SDMs demonstrated that Pleistocene climatic fluctuations had different impacts on both population sizes and the extent of suitable habitat of different populations, possibly reflecting the fact that the past climatic conditions in the Hengduan Mountains differed from those in southern China. In the following, we will discuss these findings in more detail.

At range-wide level, seven deeply divergent genetic lineages with parapatric distributions were discernable with phylogenetic, ordination and Bayesian clustering approaches. Although, we could have split W1 cluster into two further clusters from the PCA and the SNAPP tree, it would not have profoundly altered the results. These two sampling locations include only 20 individuals, which is roughly 5.7% of the data. The deep genetic divergence among different *H. annectans* lineages may be explained by two non-mutually exclusive biotic factors. First, it may be attributable to the limited dispersal capability of *H. annectans*. Geographic distance explained 23% of the variance in genetic differentiation, and a clear and strong pattern of isolation-by-distance was observed across most of the study populations. Second, the strong genetic structuring could be a result of the species-specific habitat requirements. The occurrence of *H. annectans* in China is restricted to mountain forests in low-to-

medium elevations (Fei *et al.* 2009), which suggests that dispersal over high elevation mountain ranges is unlikely (see Figure 1). Consistent with this hypothesis, we found that elevation explained over 14% of genetic differentiation of *H. annectans*.

We estimated that the seven genetic lineages of *H. annectans* diverged in the Pliocene (2.32–5.23 Mya). This timing matches the rapid and intense uplift of the QTP and adjacent Southwest China: the oldest uplifting starting about 8 Mya (Harrison *et al*. 1992) and was followed by several mid-Pliocene events (Cui *et al*. 1996; Sun *et al*. 2011). Each phase of these geological movements likely generated barriers to gene flow leading to strong genetic differentiation among populations and lineages in the southern mountains of China. For instance, the inferred divergence time of the N1 and N2 lineages at 4.4 Mya implicates the rise of Daxiangling Mountains as a probable driver of divergence. The intense uplifting of the mountain range in the western Sichuan basin, including the Daxiangling Mountains, occurred after the Miocene and reached peak elevation shortly before the Late Pliocene (Sun *et al*. 2011). Similarly, lineages W and N located in the opposite sides of the Hengduan Mountains likely diverged as a direct consequence of the uplift of the Hengduan Mountains, including Shaluli Mountains and Daxue Mountains located in this area. Further evidence for strongly restricted gene flow across the Hengduan Mountains was provided by our isolation-by-distance (IBD) analyses. We found that the pattern of IBD was weak (albeit statistically significant) when considering all populations, but became much stronger after removing populations from the high altitude lineages (W1, N1 and N2) located at opposite sides of the Hengduan Mountains. Similar uplift-driven diversification in southern China have also been reported from earlier studies of plants (Xing & Ree 2017) and animals (Yan *et al.* 2013; Li *et al.* 2015b). Thus, the complex geological events leading to habitat fragmentation and barriers to gene flow appear to be responsible for the high levels of intra- and interspecific diversity in the southern mountains of China.

Pleistocene climatic fluctuations have had strong influence on demographic processes and distribution of taxa in Europe and North America (Hewitt 2000; Hewitt 2004). In southern China, the role of

paleoclimatic fluctuations on demographic processes and distribution of taxa have remained controversial. Some studies have reported that Pleistocene climatic oscillations have shaped current patterns of genetic variation in various taxa (Ye *et al.* 2014; Li *et al*. 2018), whereas other studies have produced contradictory evidence (Wang & Ge 2006; Gao *et al*. 2011; Yan *et al*. 2013). One possible reason for these opposing results could be methodological: estimation of demographic history with limited genetic data may lead to biased inferences (Avise 2000; Ballard & Whitlock 2004; Guo *et al.* 2019). Here we were able to accurately reconstruct temporal changes in effective population size of the different lineages: these results suggest a heterogeneous effect of paleoclimatic factors on the demographic history of different *H. annectans* lineages. We found fluctuations in historical effective population sizes that followed Pleistocene climatic cycles, as well as changes in the extent and distribution of potential habitat for this species from the LIG to the present. These patterns highlight this species' sensitivity to temperature changes, in accordance with our findings that two climatic factors (BIO1 and BIO2) were strong predictors of the degree of genetic differentiation among different lineages.

The species' response to paleoclimatic changes was geographically heterogeneous. Populations distributed around the Hengduan Mountains (lineages N1, N2 and W1), displayed an obvious bottleneck following the LGM or Mid-Holocene, and had the lowest heterozygosities of all studied populations. The results from species distribution models suggest that the N1 lineage experienced severe range contraction and that the southern distribution of the N2 lineage shrank from LGM to present. Such pattern is compatible with two indistinguishable demographic scenarios: the invasion of a completely novel habitat following the retreat of an ice sheet or extinction/re-colonization. In contrast, most of the populations from lineages W2, E, C1 and C2 experienced moderate and gradual declines following the LGM or Mid-Holocene, as well as slight range contradictions (Figure 5). Such incongruence in the past demographic histories of different lineages possibly reflects differences in habitat availability. Populations in the Hengduan Mountains occupy a relatively smaller effective habitat area (Figure 1) making them more sensitive to environmental changes in their geographic range as smaller populations

have higher risk of decline or extinction (Green 2003). This was particularly relevant during the LGM – the period with dry conditions, low temperatures and extended ice sheets (Gasse 2000; Clark *et al.* 2009). The fastSTRUCTURE and TREEMIX analyses identified admixture events between the C and E, as well as the C and W2 lineages, which could be the result of homogeneous environmental condition in the Yunnan-Guizhou Plateau and Wuling Mountains. Gene flow likely increased genetic variation within these lineages, as these lineages show the highest observed and expected heterozygosities. On the other hand, low levels of contemporary and historical gene flow divergence between C with N1 and N2 lineages, in spite of geographical proximity, were associated with moderate ecological niche divergence between the C and the N1 and N2 lineages (Schoener's D = 0.230, 0.377 and Warren's I = 0.511, 0.653, respectively, Table 2).

A clear and strong IBD was observed across most of the study populations, suggesting that there is – or has been – some degree of gene flow between different lineages; this conclusion was also supported by the TREEMIX results. However, the N1 and N2 lineages were an exception. These are the two lineages that are geographically closest (38 km), and yet displayed extremely high genetic divergence (F_{ST} = 0.909), suggesting the absence of gene flow. Such divergence is likely the result of geological barriers (in this case: Daxiangling Mountains) and strong genetic drift. Interestingly, these lineages also showed moderate niche divergence (Schoener's D = 0.337 and Warren's I = 0.634), suggesting that they may be also ecologically divergent. However, the current distribution range of the N1 lineage may not reflect its past distribution. As shown by the SDM results, the habitat suitable for the N1 lineage around the northwest corner of Sichuan Basin contracted gradually from the LIG to the present. This interpretation is also supported by the observation that the populations in the N1 and N2 lineages were found to have experienced population size bottlenecks at different time periods.

Apart from the deep genetic divergences revealed by both phylogenetic analyses and high pairwise F_{ST} values, the seven major lineages exhibited clear differentiation in phenotypic traits. Although differentiation in the mean values of all studies traits was significant, DRIFTSEL analyses suggest that

only divergence in the number of black spots was likely driven by selection. Since the data utilized in these analyses came from wild-collected individuals, rather than from a common garden experiment, the results should be interpreted with caution (Karhunen *et al*. 2013; Leinonen *et al*. 2013). Namely, the possibility of environmentally induced differences cannot be excluded as an alternative explanation for divergence in the number of black spots. Regardless of whether this divergence was genetic, environmental, or due to their combined effects, it is noteworthy that the magnitude of phenotypic differentiation (P_{ST}) exceeded, on average, that of neutral genetic differentiation (F_{ST}). Since the neutral expectation is that $P_{ST} \approx F_{ST}$, any deviation from this calls for an explanation. Similarly, the fact that phenotypic differentiation was a negative function of neutral differentiation is noteworthy. The mechanistic explanation for this negative correlation is that some of the most genetically divergent populations were phenotypically the least diverged (e.g. lineages N1, N2 and E in Figure 2). Without common garden data, we cannot conclusively establish an ultimate explanation for the observed patterns. Nevertheless, we suspect that the divergence in number of black spots may be related to different sexual selection regimes in different lineages, as observed in other systems (Endler 1983; Reynolds & Fitzpatrick 2007; Rudh *et al*. 2007). Hence, an interesting avenue for future studies would be to investigate whether the phenotypic differentiation in number of black spots might act as pre-zygotic isolation mechanism between the seven genetically divergent lineages.

Conclusions

In conclusion, our genome-wide analyses of *H. annectans* revealed seven highly differentiated genetic lineages, which also show clear phenotypic differences likely attributable to the action of natural selection. The estimated divergence times for these lineages closely align with the timing of the uplifting of the QTP and adjacent southwest China, suggesting that past geological events played a major role in shaping the distribution of genetic diversity within this species complex. Populations living in different areas displayed different demographic dynamics in response to Pleistocene and Holocene climatic changes. This is expected, because of the geographic and temporal variation in the climatic conditions experienced in different areas. As such, this study provides an example of how the combined effects of

geomorphological and climatic factors have shaped the distribution of genetic variation in a widely distributed species. The results highlight how geological events and topographic features play predominant roles as drivers of lineage differentiation, and that climatic fluctuations contribute to re-shaping the distribution of genetic variability.

Acknowledgments

We thank Chengliang Li, Wanxin Du, Jibing Liu, Youming Zhang and Xiaoran Zhu help with sample collection. Xueyun Feng and Bohao Fang kindly helped with solving analytical problems. We thank Jacquelin De Faveri for a linguistic check on an earlier version of this manuscript. We are also thankful for the computing resource support from CSC - the Finnish IT Center for Science Ltd administered by the Ministry of Education and Culture, Finland. Our research was supported by grants from the Biodiversity Survey, Monitoring and Assessment Project of Ministry of ecology and environment, China (2019–2023) and Academy of Finland (No. 129662, 134728 and 218343 to Juha Meriläand No. 316294 to Paolo Momigliano).

References

Avise JC, Arnold J, Ball RM *et al.* (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 489-522.

Avise JC (1994) *Molecular Markers, Natural History and Evolution*. Chapman & Hall, New York.

Avise JC (2000) *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, Massachusetts.

Avise JC (2009) Phylogeography: retrospect and prospect. *Journal of Biogeography*, **36**, 3-15.

Ballard JW, Whitlock MC (2004) The incomplete natural history of mitochondria. *Molecular Ecology*, **13**, 729-744.

Bossuyt F, Milinkovitch MC (2001) Amphibians as indicators of early tertiary "out-of-India" dispersal of vertebrates. *Science*, **292**, 93-95.

Bouckaert R, Heled J, Kühnert D *et al.* (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, **10**, e1003537.

Brommer JE (2011) Whither P_{ST}? The approximation of Q_{ST} by P_{ST} in evolutionary and conservation biology. *Journal of Evolutionary Biology*, **24**, 1160-1168.

Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, RoyChoudhury A (2012) Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution*, **29**, 1917-1932.

Chaves JA, Weir JT, Smith TB (2011) Diversification in *Adelomyia* hummingbirds follows Andean uplift. *Molecular Ecology*, **20**, 4564-4576.

Che J, Zhou WW, Hu JS *et al.* (2010) Spiny frogs (Paini) illuminate the history of the Himalayan region and Southeast Asia. *Proceedings of the National Academy of Sciences USA*, **107**, 13765-13770.

Clark PU, Dyke AS, Shakun JD *et al.* (2009) The last glacial maximum. *Science*, **325**, 710-714.

Clevenger J, Chavarro C, Pearl SA, Ozias-Akins P, Jackson SA (2015) Single nucleotide polymorphism identification in polyploids: a review, example, and recommendations. *Molecular Plant*, **8**, 831-846.

Cobos ME, Peterson AT, Barve N *et al.* (2019) kuenm: an R package for detailed development of ecological niche models using Maxent. *PeerJ*, **7**, e6281.

Cui ZJ, Gao QZ, Liu GN *et al.* (1996) Planation surfaces, palaeokarst and uplift of Xizang (Tibet) Plateau. *Science in China Series D: Earth Sciences*, **39**, 391-400.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.
Dufresnes C, Strachinis I, Suriadna N *et al*. (2019) Phylogeography of a cryptic speciation continuum in Eurasian spadefoot toads (*Pelobates*). *Molecular Ecology*, **28**, 3257-3270.
Endler JA (1983) Natural and sexual selection on color patterns in poeciliid fishes. *Environmental Biology of Fishes*, **9**, 173-190.
Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564-567.
Fei L, Hu SQ, Ye CY *et al.* (2009) *Fauna Sinica, Amphibia: Vol 2. Anura, Hylidae*. Science Press, Beijing.
Feng SH, Fang Q, Barnett R *et al*. (2019) The genomic footprints of the fall and recovery of the crested ibis. *Current Biology*, **29**, 340-349.
Francis RM (2017) Pophelper: An R package and web app to analyse and visualize population structure. *Molecular Ecology Resources*, **17**, 27-32.
Gao B, Yu LJ, Qu YH *et al.* (2011) An unstructured phylogeographic pattern with extensive gene flow in an endemic bird of South China: Collared finchbill (*Spizixos semitorques*). *International Journal of Molecular Sciences*, **12**, 3635-3647.
Gasse F (2000) Hydrological changes in the African tropics since the Last Glacial Maximum. *Quaternary Science Reviews*, **19**, 189-211.
Graham CH, Ron SR, Santos JC, Schneider CJ, Moritz C (2004) Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. *Evolution*, **58**, 1781-1793.
Green DM (2003) The ecology of extinction: population fluctuation and decline in amphibians. *Biological Conservation*, **111**, 331-343.
Guo BC, Fang BH, Shikano T, Momigliano P, Wang C, Kravchenko A, Meril äJ (2019) A phylogenomic perspective on diversity, hybridization and evolutionary affinities in the stickleback genus *Pungitius*. *Molecular Ecology*, **28**, 4046-4064.
Hanania U, Velcheva M, Sahar N, Perl A (2004) An improved method for isolating high-quality DNA from *Vitis vinifera* nuclei. *Plant Molecular Biology Reporter*, **22**, 173-177.
Harrison TM, Copeland P, Kidd WS, Yin A (1992) Raising Tibet. *Science*, **255**, 1663-1670.
Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247-276.
Hewitt GM (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907-913.
Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **359**, 183-195.
Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965-1978.
Jiang WS, Qiu Y, Pan XF *et al*. (2018) Genome assembly for a Yunnan-Guizhou Plateau "3E" fish, *Anabarilius grahami* (Regan), and its evolutionary and genetic applications. *Frontiers in Genetics*, **9**, 614.
Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403-1405.
Ju LX, Wang HJ, Jiang DB (2007) Simulation of the Last Glacial Maximum climate over East Asia with a regional climate model nested in a general circulation model. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **248**, 376-390.
Karhunen M, Meril äJ, Leinonen T *et al.* (2013) Driftsel: an R package for detecting signals of natural selection in quantitative traits. *Molecular Ecology Resources*, **13**, 746-754.
Karhunen M, Ovaskainen O, Herczeg G *et al.* (2014) Bringing habitat information into statistical tests of local adaptation in quantitative traits: A case study of nine-spined sticklebacks. *Evolution*, **68**, 559-568.
Kent WJ (2002) BLAT-the BLAST-like alignment tool. *Genome Research*, **12**, 656-664.
Kierepka EM, Latch EK (2014) Performance of partial statistics in individual-based landscape genetics. *Molecular Ecology Resources*, **15**, 512-525.

Korneliussen TS, Albrechtsen A, Nielsen R (2014) ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics*, **15**, 356.
Kozak KH, Wiens JJ (2010) Niche conservatism drives elevational diversity patterns in Appalachian salamanders. *American Naturalist*, **176**, 40-54.
Kumar R, Kumar V (2018) A review of phylogeography: biotic and abiotic factors. *Geology, Ecology, and Landscapes*, **2**, 268-274.
Kumar S, Stecher G, Li M *et al.* (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, **35**, 1547-1549.
Leinonen T, Scott McCairns RJ, O'Hara RB, Merilä J (2013) Q_{ST}-F_{ST} comparisons: evolutionary and ecological insights from genomic heterogeneity. *Nature Reviews Genetics*, **14**, 179-190.
Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754-1760.
Li H, Handsaker B, Wysoker A *et al.* (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078-2079.
Li JT, Wang JS, Nian HH *et al.* (2015a) Amphibians crossing the bering land bridge: evidence from holarctic treefrogs (*Hyla*, Hylidae, Anura). *Molecular Phylogenetics and Evolution*, **87**, 80-90.
Li J, Zhao M, Wei SC *et al.* (2015b) Geologic events coupled with Pleistocene climatic oscillations drove genetic variation of Omei treefrog (*Rhacophorus omeimontis*) in southern China. *BMC Evolutionary Biology*, **15**, 289.
Li J, Wei SC, Hu ML *et al.* (2018) Reflection of paleoclimate oscillations and tectonic events in the phylogeography of moustache toads in southern China. *Journal of Zoology*, **305**, 17-26.
Liao WB, Lu X (2010) Age structure and body size of the Chuanxi Tree Frog *Hyla annectans chuanxiensis* from two different elevations in Sichuan (China). *Zoologischer Anzeiger-A Journal of Comparative Zoology*, **248**, 255-263.
Liu K (1988) Quaternary history of the temperate forests of China. *Quaternary Science Reviews*, **7**, 1-20.
Liu XM, Fu YX (2015) Exploring population size changes using SNP frequency spectra. *Nature Genetics*, **47**, 555-559.
McCormack JE, Hird SM, Zellmer AJ *et al.* (2013) Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution*, **66**, 526-538.
McKenna A, Hanna M, Banks E *et al.* (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, **20**, 1297-1303.
Muscarella R, Galante PJ, Soley-Guardia M *et al.* (2014). ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. *Methods in Ecology and Evolution*, **5**, 1198-1205.
Newman CE, Austin CC (2016) Sequence capture and next generation sequencing of ultraconserved elements in a large genome salamander. *Molecular Ecology*, **25**, 6162-6174.
Oksanen FJ, Blanchet FG, Friendly M *et al.* (2019). *Vegan: Community ecology package*. R package version 2.5-6. https://cran.r-project.org/package=vegan
O'Rawe J, Jiang T, Sun GQ *et al.* (2013) Low concordance of multiple variant-calling pipelines: practical implications for exome and genome sequencing. *Genome Medicine*, **5**, 28.
Ovaskainen O, Karhunen M, Zheng CZ *et al.* (2011) A new method to uncover signatures of divergent and stabilizing selection in quantitative traits. *Genetics*, **189**, 621-632.
Owens HL, Campbell LP, Dornak LL *et al.* (2013) Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. *Ecological Modelling*, **263**, 10-18.
Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. *PLoS Genetics*, **2**, e190.
Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231-259.
Pickrell JK, Pritchard JK (2012) Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics*, **8**, e1002967.
Puckett EE, Park J, Combs M *et al.* (2016) Global population divergence and admixture of the brown rat (*Rattus norvegicus*). *Proceedings of the Royal Society B: Biological Sciences*, **283**, 20161762.

R Core Team (2014) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/
Raj A, Stephens M, Pritchard JK (2014) fastSTRUCTURE: Variational inference of population structure in large SNP data sets. *Genetics*, **197**, 573-589.
Rambaut A, Drummond AJ, Xie D, *et al*. (2018) Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, **67**, 901.
Reynolds RG, Fitzpatrick BM (2007) Assorted mating in poison-dart frogs based on an ecologically important trait. *Evolution*, **61**, 2253-2259.
Rudh A, Rogell B, Höglund J (2007) Non-gradual variation in color morphs of the strawberry poison frog *Dendrobates pumilio*: genetic and geographical isolation suggest a role for selection in maintaining polymorphism. *Molecular Ecology*, **16**, 4284-4294.
Sánchez-Montes G, Wang JL, Ariño AH *et al*. (2018) Mountains as barriers to gene flow in amphibians: Quantifying the differential effect of a major mountain ridge on the genetic structure of four sympatric species with different life history traits. *Journal of Biogeography*, **45**, 318-331.
Schoener TW (1968) The *Anolis* lizards of Bimini: Resource partitioning in a complex fauna. *Ecology*, **49**, 704-726.
Shen YH (1996) A new subspecies of *Hyla annectans* from Hunan, China (Anaura: Hylidae). *Zoological Research*, **18**, 177-182.
Shi YF, Ren B, Wang J, Derbyshire E (1986) Quaternary glaciation in China. *Quaternary Science Reviews*, **5**, 503-507.
Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457-462.
Sokal RR, Rohlf FJ (1981) *Biometry*. W. H. Freeman and Co, San Francisco, USA.
Stange M, Sánchez-Villagra MR, Salzburger W *et al.* (2018) Bayesian divergence-time estimation with genome-wide single-nucleotide polymorphism data of sea catfishes (Ariidae) supports Miocene closure of the Panamanian Isthmus. *Systematic Biology*, **67**, 681-699
Sun BN, Wu JY, Liu YS *et al.* (2011) Reconstructing Neogene vegetation and climates to infer tectonic uplift in western Yunnan, China. *Palaeogeography Palaeoclimatology Palaeoecology*, **304**, 328-336.
Sun XW, Liu DY, Zhang XF *et al.* (2013) SLAF-seq: an efficient method of large-scale *de novo* SNP discovery and genotyping using high-throughput sequencing. *PLoS ONE*, **8**, e58700.
Sun YB, Xiong ZJ, Xiang XY *et al.* (2015) Whole-genome sequence of the Tibetan frog *Nanorana parkeri* and the comparative evolution of tetrapod genomes. *Proceedings of the National Academy of Sciences USA*, **112**, E1257-E1262.
Wang HW, Ge S (2006) Phylogeography of the endangered *Cathaya argyrophylla* (Pinaceae) inferred from sequence variation of mitochondrial and nuclear DNA. *Molecular Ecology*, **15**, 4109-4122.
Wang GD, Zhang BL, Zhou WW *et al*. (2018) Selection and environmental adaptation along a path to speciation in the Tibetan frog *Nanorana parkeri*. *Proceedings of the National Academy of Science*s, **115**, E5056-E5065.
Ward RD, Skibinsky DOF, Woodwark M (1992) Protein heterozygosity, protein structure and taxonomic differentiation. *Evolutionary Biology*, **26**, 3-131.
Warren DL, Glor RE, Turelli M (2008) Climate niche identity versus conservatism: Quantitative approaches to niche evolution. *Evolution*, **62**, 2868-2883.
Watanabe S, Hajima T, Sudo K *et al.* (2011) MIROC-ESM 2010: Model description and basic results of CMIP5-20c3m experiments. *Geoscientific Model Development*, **4**, 845-872.
Weaver AJ, Eby M, Fanning AF, Wiebe EC (1998) Simulated influence of carbon dioxide, orbital forcing and ice sheets on the climate of the Last Glacial Maximum. *Nature*, **394**, 847-853.
Xing YW, Ree RH (2017) Uplift-driven diversification in the Hengduan Mountains, a temperate biodiversity hotspot. *Proceedings of the National Academy of Sciences USA*, **114**, E3444-E3451.
Yan F, Zhou WW, Zhao HT *et al.* (2013) Geological events play a larger role than Pleistocene climatic fluctuations in driving the genetic structure of *Quasipaa boulengeri* (Anura: Dicroglossidae). *Molecular Ecology*, **22**, 1120-1133.

Ye Z, Zhu GP, Chen PP, Zhang DL, Bu WJ (2014) Molecular data and ecological niche modelling reveal the Pleistocene history of a semi-aquatic bug (*Microvelia douglasi douglasi*) in East Asia. *Molecular Ecology*, **23**, 3080-3096.
Zeisset I, Beebee TJC (2008) Amphibian phylogeography: a model for understanding historical aspects of species distributions. *Heredity*, **101**, 109-119.
Zhang RZ (1999) *Zoogeography of China*. Science Press, Beijing, China.
Zhao J, Shi Y, Wang J (2011) Comparison between Quaternary glaciations in China and the Marine Oxygen Isotope Stage (MIS): an improved schema. *Acta Geographica Sinica*, **66**, 867-884.
Zhao SC, Zheng PP, Dong SS *et al*. (2013) Whole-genome sequencing of giant pandas provides insights into demographic history and local adaptation. *Nature Genetics*, **45**, 67.
Zhao J, Perez MBM, Hu J, Fernandez MGS (2016) Genome-wide association study for nine plant architecture traits in Sorghum. *Plant Genome*, **9**, 2.
Zheng B, Xu Q, Shen Y (2002) The relationship between climate change and Quaternary glacial cycles on the Qinghai-Tibetan Plateau: review and speculation. *Quaternary International*, **97**, 93-101.
Zhou WJ, Yu XF, Jull AT *et al.* (2004). High-resolution evidence from southern China of an early Holocene optimum and a mid-Holocene dry event during the past 18 000 years. *Quaternary Research*, **62**, 39-48.
Zhou XM, Meng XH, Liu ZJ *et al*. (2016) Population genomics reveals low genetic diversity and adaptation to hypoxia in snub-nosed monkeys. *Molecular Biology and Evolution*, **33**, 2670-2681.

Data Accessibility

Aligned datasets, all ascii files from the SDM along with the all morphological data deposited in Dryad (https://datadryad.org/stash/share/TvAr1XyODIzwAfP6zV4_7sow-2ucxSr85T9Dfo4RcxY). Sampling locations and additional individuals included in SDM analyses are uploaded as online Supporting Information. All raw sequence data are uploaded to Genbank's Short Read Archive (Accession nos. SRR12349892–SRR12350250).

Author contributions

H.W., J.M.: designed the research and contributed to draft the manuscript. C.F.: contributed to sampling of materials. S.W., Z.L. and P.M.: analyzed the data. S.W., P.M. and J.M. interpreted the results. S.W wrote the manuscript with significant contributions from other authors. All authors read and approved the final manuscript.

Conflicts of interest

Authors declare no conflict of interests

Tables and Figures

Table 1. Results of dbRDA analysis testing for the effect of geographic (GEO1-3, Elevation) and environmental (BIO1, BIO2, BIO12, BIO14) on the degree of genetic differentiation (as measured by F_{ST}) among 35 populations of *H. annectans*. See text for explanation of GEO and BIO variables.

	% of variance explained	*d.f.*	*F*	*P*
Geography	0.37			
Elevation	0.14	1	5.34	0.001
GEO1	0.13	1	4.82	0.001
GEO2	0.08	1	3.04	0.001
GEO3	0.02	1	0.81	0.597
Environment	0.41			
BIO1	0.06	1	2.12	0.033
BIO2	0.13	1	5.01	0.002
BIO12	0.08	1	2.73	0.004
BIO14	0.14	1	5.39	0.001
Residual	0.22	26		

Table 2. Summary of niche divergence comparisons among five *H. annectans* lineages using Schonner's *D* (above diagonal) and Warren's *I* (below diagonal).

Lineages		N1	N2	W	C	E
				Schoener's *D*		
N1		1.000	0.337	0.542	0.230	0.154
N2		0.634	1.000	0.394	0.377	0.118
W	Warren's *I*	0.831	0.679	1.000	0.456	0.238
C		0.511	0.653	0.774	1.000	0.438
E		0.370	0.296	0.488	0.745	1.000

Table 3. Results of Kruskal-Wallis and subsequent *post hoc* (Dunn's multiple comparisons test) tests for differences in mean number of black spots among the seven main *H. annectans* lineages.

Lineage	N	Mean rank	Kruskal-Wallis test value	*P*	Dunn's multiple comparison test					
					C1	C2	N1	N2	W1	W2
E	100	281.30	258.22	<0.001	*	*	*	*	*	*
C1	40	202.13				*	ns	ns	*	*

C2	30	112.25	ns	ns	ns	ns
N1	29	187.03		ns	*	ns
N2	20	158.50			*	ns
W1	90	67.62				ns
W2	30	112.25				

* $P < 0.05$

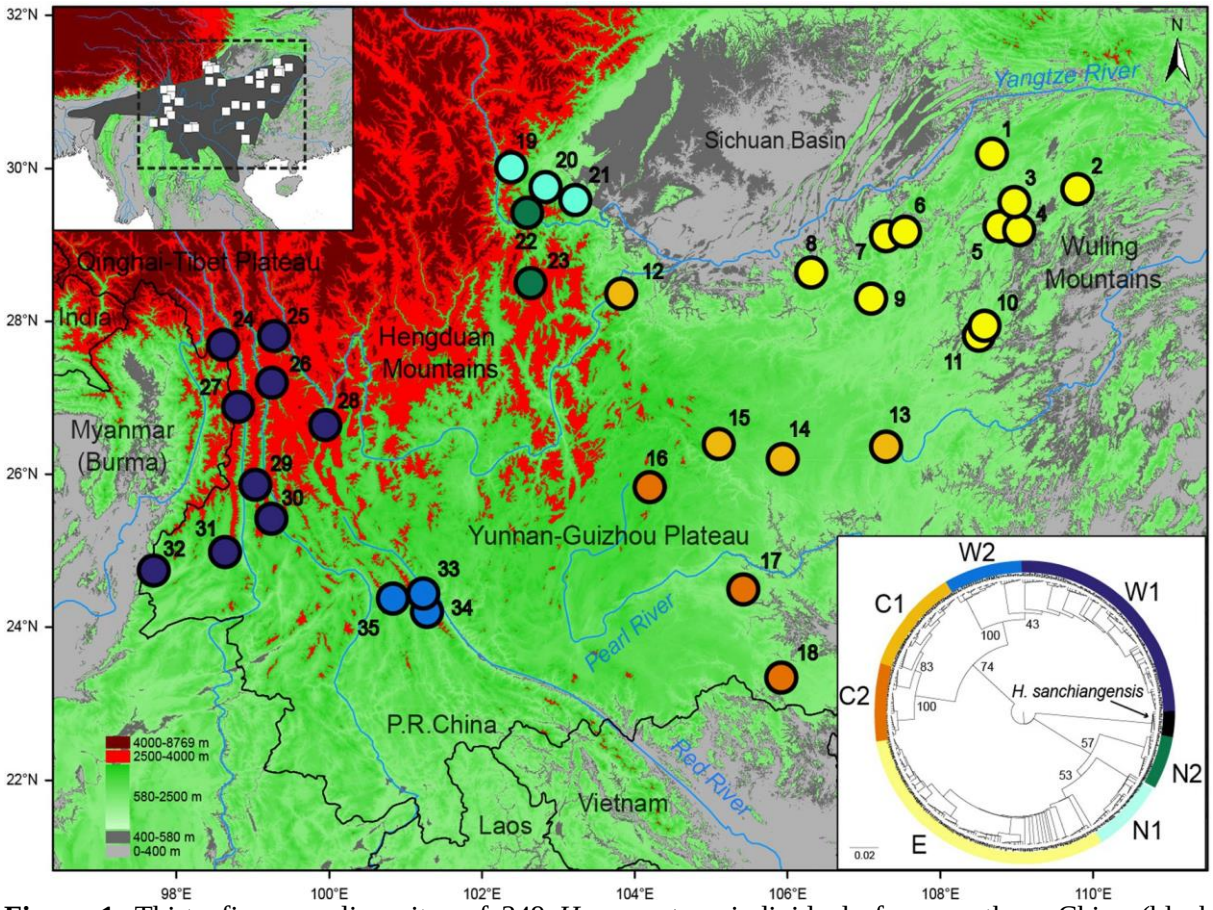

Figure 1. Thirty five sampling sites of 349 *H. annectans* individuals from southern China (black symbols). The grey shade in the left corner insert indicates the entire species distribution range, downloaded from the IUCN website (http://www.iucnredlist.org/). The primary mountain systems and basins are indicated, and the major rivers are depicted with blue lines. The gradient of color on the map represents different elevations, the areas exceeding values suitable for *H. annectans* are shown in red (high elevation) and grey (low elevation). Populations are numbered as in Table S1. The phylogenetic tree on the right corner insert is a NJ tree constructed in MEGA X, branch labels represent bootstrap support value. Seven genetic lineages are indicated with different colors, which are also used in all other figures.

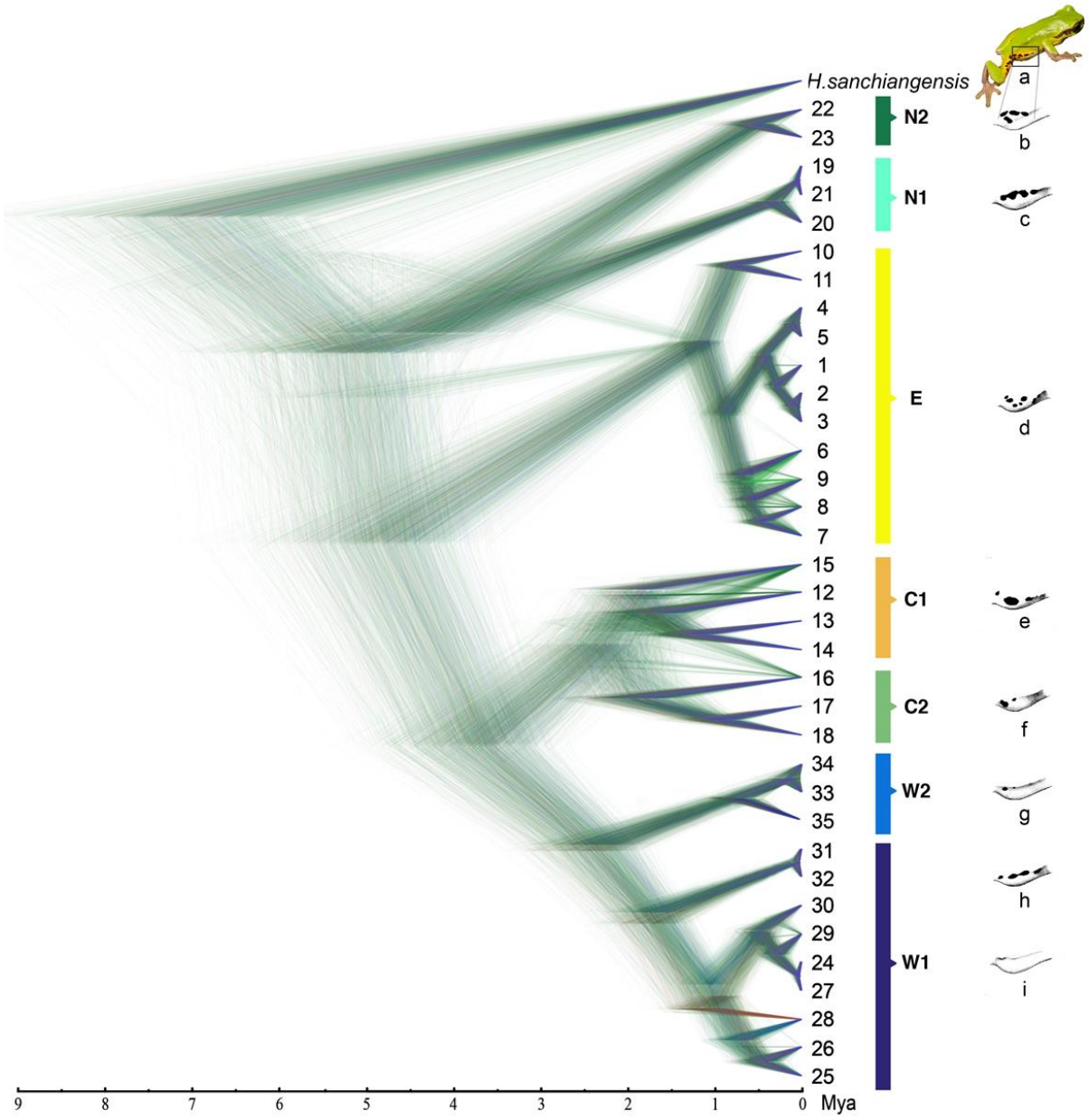

Figure 2. Phylogenetic relationships and phenotypic variation among *H. annectans* lineages. The SNAPP tree was obtained using *H. sanchigangensis* as an outgroup. Divergence times are shown on the horizontal axis. Colored boxes highlight the seven lineages consistent with genetic clustering analyses. The number of black spots on flanks is given on the right: a, *H. annectans* (this individual was not included in this study); b, individual "fp" from site 22 and genetic cluster N2 (4 black spots); c, individual "fg" from site 19 and genetic cluster N1 (4 black spots); d, individual "bf" from site 11 and genetic cluster E with (7 black spots); e, individual from site 14 and genetic cluster C1 (3 black spots); f, individual "ku" from site 18 and genetic cluster C2 (2 black spots); g, individual "lq" from site 35 and genetic cluster W2 (2 black spots); h, individual "nk" from site 31 and genetic cluster W1 (4 black spots); i, individual "jg" in site 27 and genetic cluster W1 (no black spots).

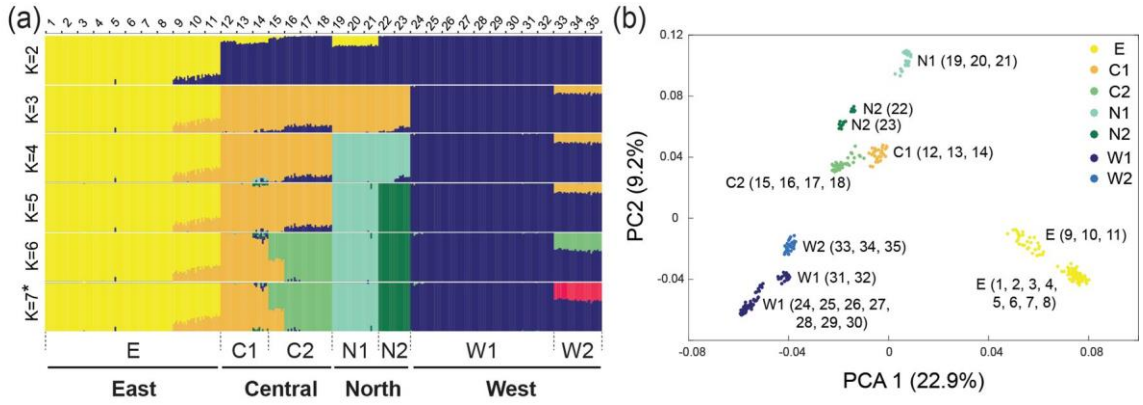

Figure 3. Inferred genetic structure of *H. annectans* populations according to (a) Bayesian cluster analysis using fastSTRUCTURE from K = 2 to K = 7 based on the 8,420 SNP dataset, and (b) PCA based on the 8,420 SNP dataset. In (a), codes above and below the plot refer to population and cluster identifiers, respectively. Different clusters are indicated with different colors. *denotes the optimal K value. In (b) the different colors of individual data points are coded according to their cluster identities.

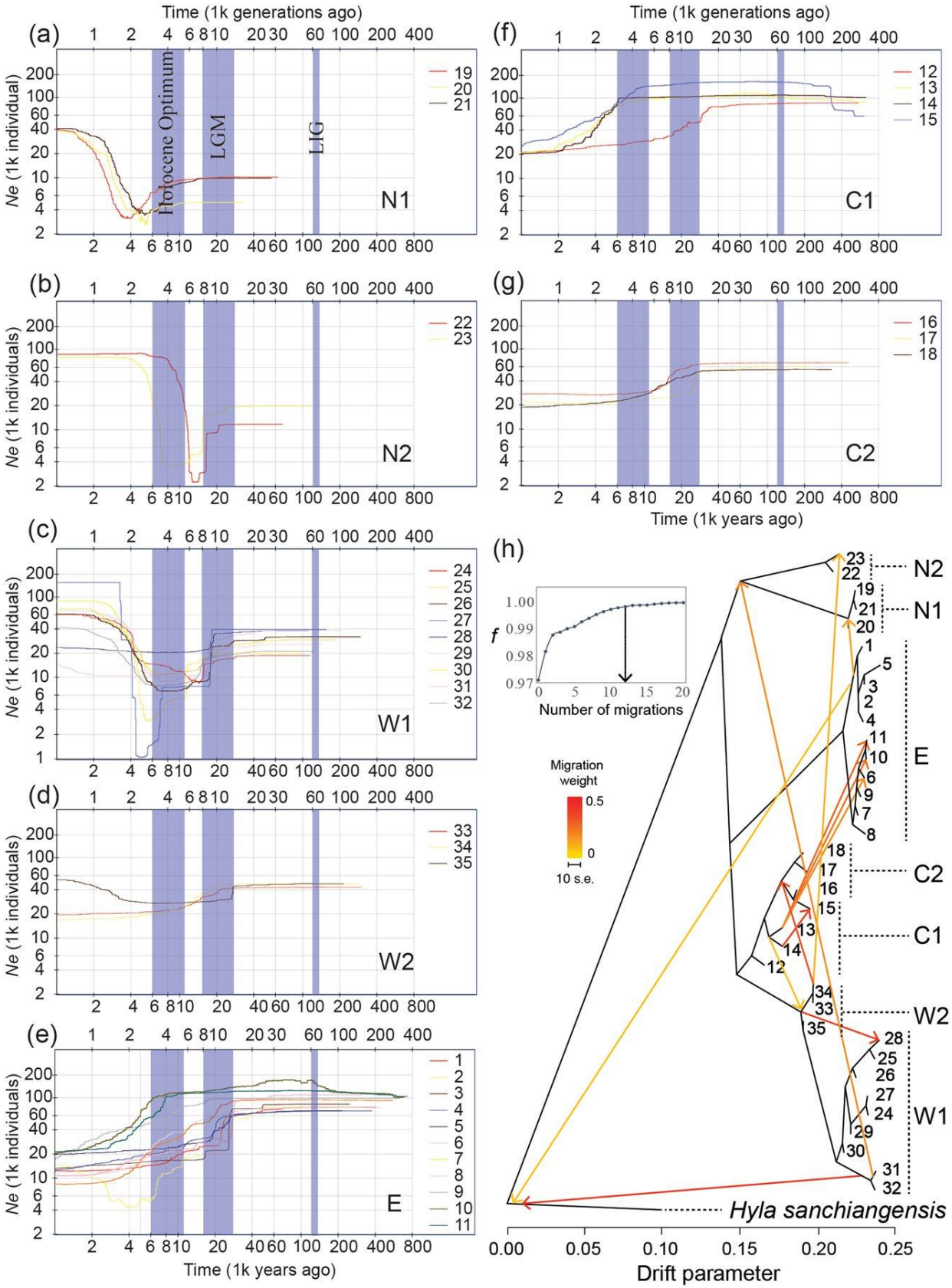

Figure 4. Demographic history of *H. annectans* lineages. The x-axis indicates time, and the y-axis indicates N_e. Different colored lines in the plot depict different populations within a given lineage. (a) to (g) are STAIRWAY PLOTS for populations of the N1, N2, W1, W2, E, C1 and C2 lineages,

respectively. The blue shaded areas mark the Holocene Optimum (6–11 Kya, constrained by Zhou *et al*. 2004), LGM (16–28 kya, constrained by Zhao *et al*. 2011) and LIG (120–140 kya, WorldClim: http://www.worldclim.org/) periods. (h) Migration among lineages inferred by TREEMIX. The migration weight indicates the proportion of ancestry derived from the migration edge.

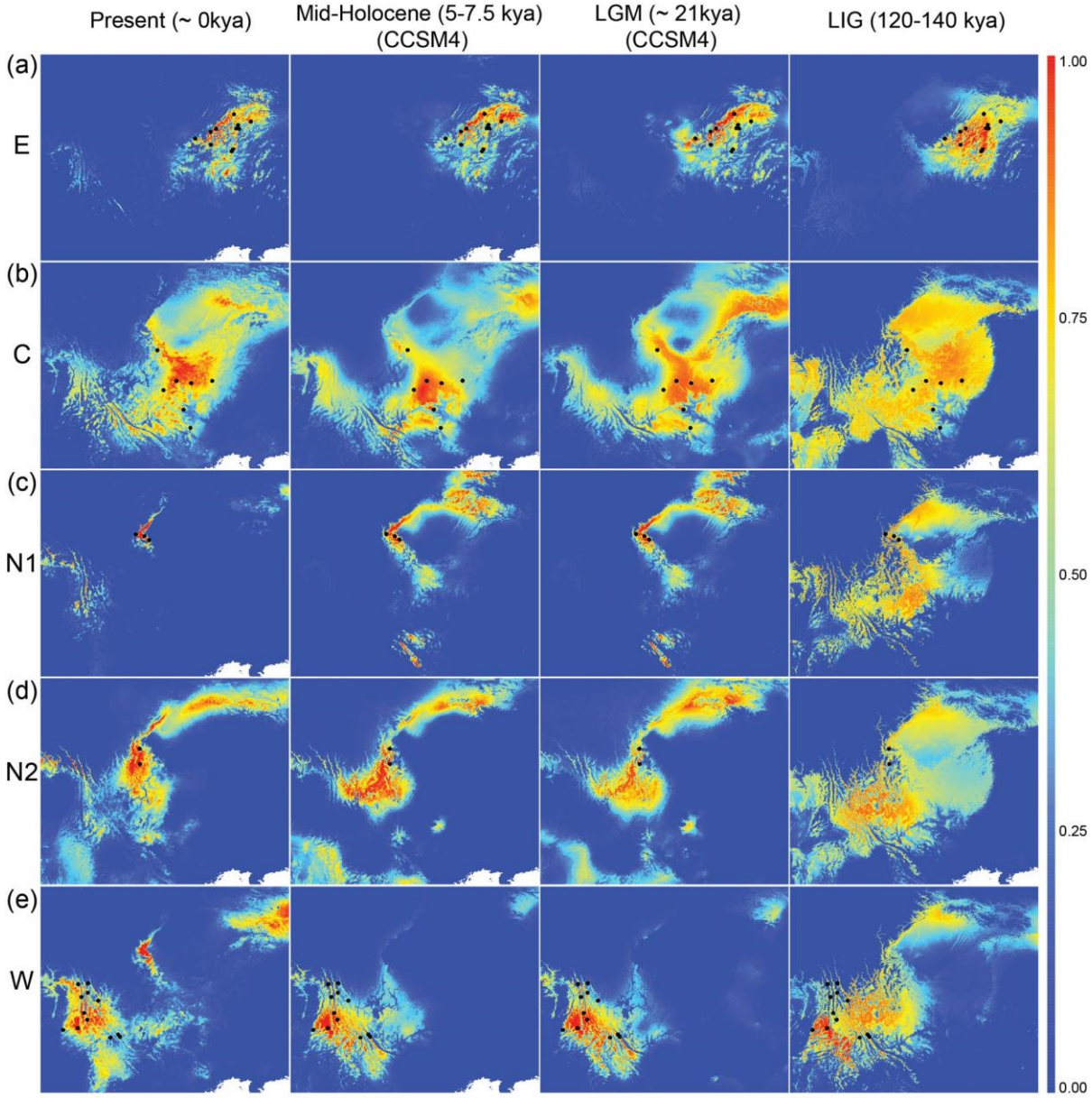

Figure 5. Species distribution models for *H. annectans* for present and historical (Mid Holocene (5–7.5 kya), Last Glacial Maximum [LGM], 21 kya and the Last Interglacial [LIG], 120–140 kya) times. Warmer colors indicate higher probability of occurrence as predicted by MAXENT (Phillips *et al.* 2006). Black symbols depict sampling sites used in this study. (a) to (e) are SDMs for populations of the E, C, N1, N2, and W lineages, respectively.

857 **Supporting information**

858 Additional supporting information may be found in the online version of this article:

859 **Table S1**. Sampling site data, gender and distribution cluster for genetic samples of *H. annenctans* used
860 in this study.

861 **Table S2.** Summary of each dataset used for the respective analysis.

862 **Table S3.** GPS points used for building species distribution models for *H. annectans*.

863 **Table S4.** Bioclimatic variable selection for species distribution modelling based on PCA analysis.

864 **Table S5.** Summary of SLAF data collected in the final assembly of 349 samples of *H. annectans*.

865 **Table S6.** Summary statistics of average observed heterozygosity (*Ho*), average expected
866 heterozygosity (*He*) and pairwise F_{ST} values between the 35 populations of *H. annectans*. All F_{ST} values
867 are statistically significant ($P < 0.05$).

868 **Table S7.** Best fit MAXENT model based upon delta AICc from ENMeval.

869 **Table S8.** AUC values for each model.

870 **Table S9.** Pairwise *F*-values of morphological features (snout-vent length [SVL] and weight) of 339
871 individuals from the seven main genetic lineages of *H. annectans*. Data were analyzed with Welch
872 ANOVA using a Games-Howell *post hoc* test.

873 **Figure S1.** SNAPP tree with the 95% posterior distributions of the time calibration.

874 **Figure S2.** Results from fastSTRUCTURE clustering analyses using the simple models. (a) Prediction
875 error from five-fold cross-validation. The lowest value of CV error is when $K = 7$. (b) The marginal
876 likelihood at increasing number of *K*. The marginal likelihood is maximized when $K = 7$.

877 **Figure S3.** Genetic isolation by distance.

878 **Figure S4.** Variance partitioning results of dbRDA analyses.

879 **Figure S5.** Delta AICc values of all models compared in ENMeval.

880 **Figure S6.** SDMs for other models.

881 **Figure S7.** Mobility-oriented parity analysis with during three periods (HM, LGM, and LIG).

Supporting Information for:

The roles of climate, geography and natural selection as drivers of genetic and phenotypic differentiation in a widespread amphibian *Hyla annectans* (Anura: Hylidae)

Shichao Wei[1], Zitong Li[2], Paolo Momigliano[2], Chao Fu[1], Hua Wu[1]*, & Juha Merilä[2]

[1]*Institute of Evolution and Ecology, School of Life Sciences, Central China Normal University, 152 Luoyulu, Hongshan District, 430079 Wuhan, China*

[2]*Ecological Genetics Research Unit, Organismal and Evolutionary Biology Research Programme, Faculty of Biological and Environmental Sciences, University of Helsinki, P.O. Box 65, FI-00014 Helsinki, Finland*

***Corresponding author**: Hua Wu, wuhua@mail.ccnu.edu.cn*

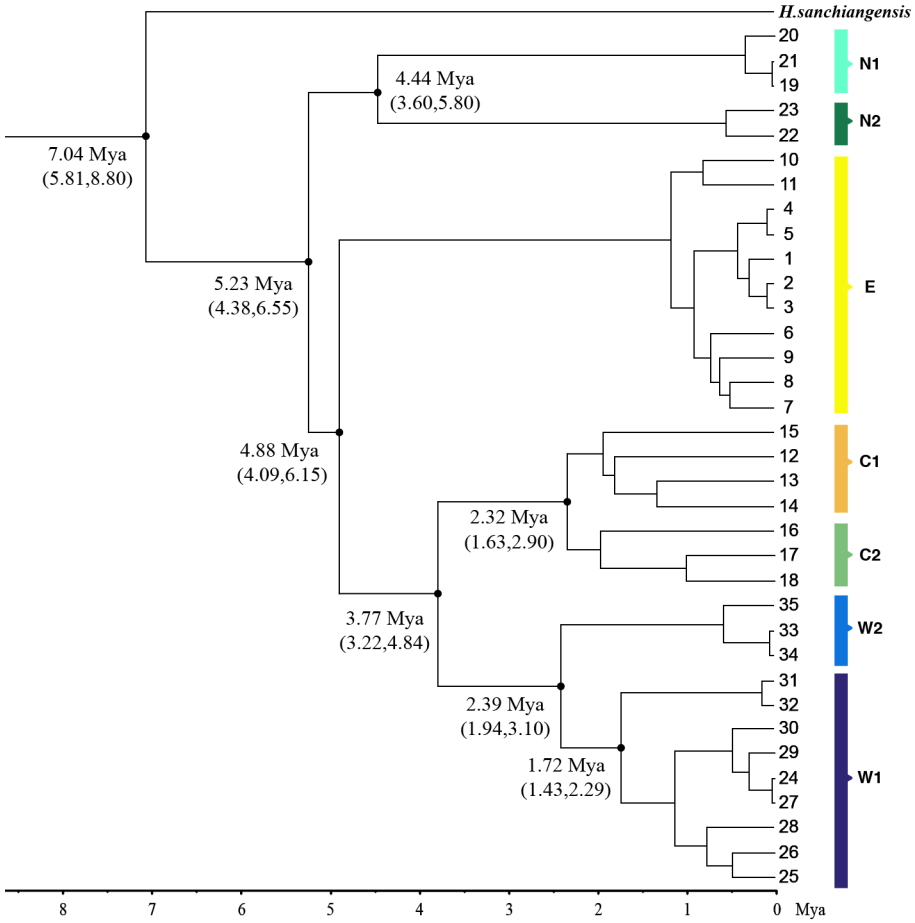

Figure S1. SNAPP tree illustrating phylogenetic relationships among *H. annectans* using *H. sanchigangensis* as an outgroup. The lineage divergence times were calibrated and the numbers in the brackets represent the 95% posterior distributions of the estimates. Colored boxes highlight the seven lineages consistent with genetic clustering analyses.

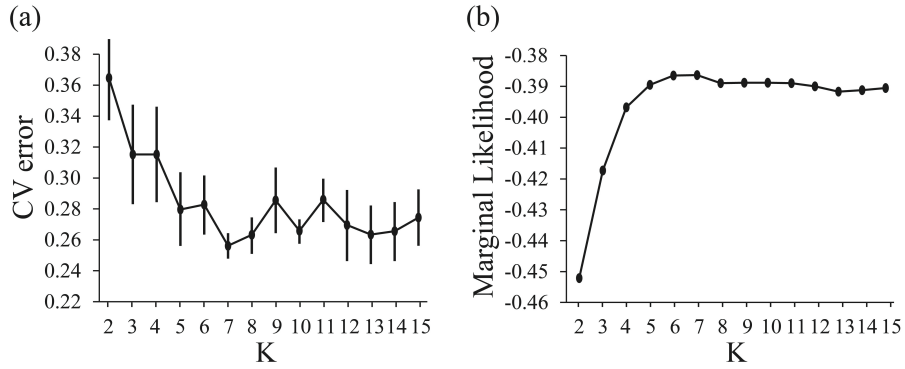

Figure S2. Results from fastSTRUCTURE clustering analyses using the simple model. (a) Prediction error from fivefold cross-validation (CV) for the fastSTRUCTURE analyses, the lowest value of CV error is when $K = 7$. (b) The marginal likelihood at increasing number of K. The marginal likelihood is maximized when $K = 7$.

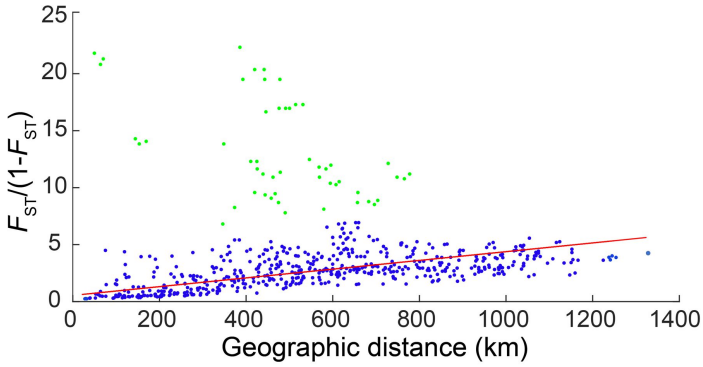

Figure S3. Correlation between pairwise genetic differentiation among populations (Slatkin's linearized F_{ST}) and the geographic distance separating populations. The green dots indicate comparisons of populations in clusters N1, N2 and W1, and were excluded from the other correlation analysis. The red line is the slope of linear regression (excluding the green data points) given solely for illustrative purposes.

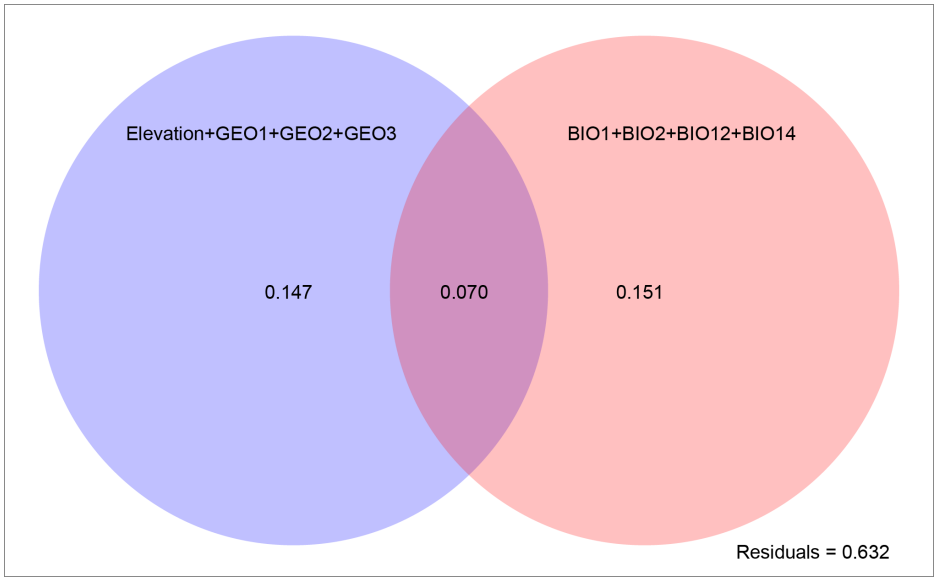

Figure S4. Variance partitioning results of dbRDA analyses. Geographic variables comprises four geographic vectors (Elevation, GEO1, GEO2, GEO3). Environmental variables comprise four environmental variables: BIO1, BIO2, BIO12, BIO14.

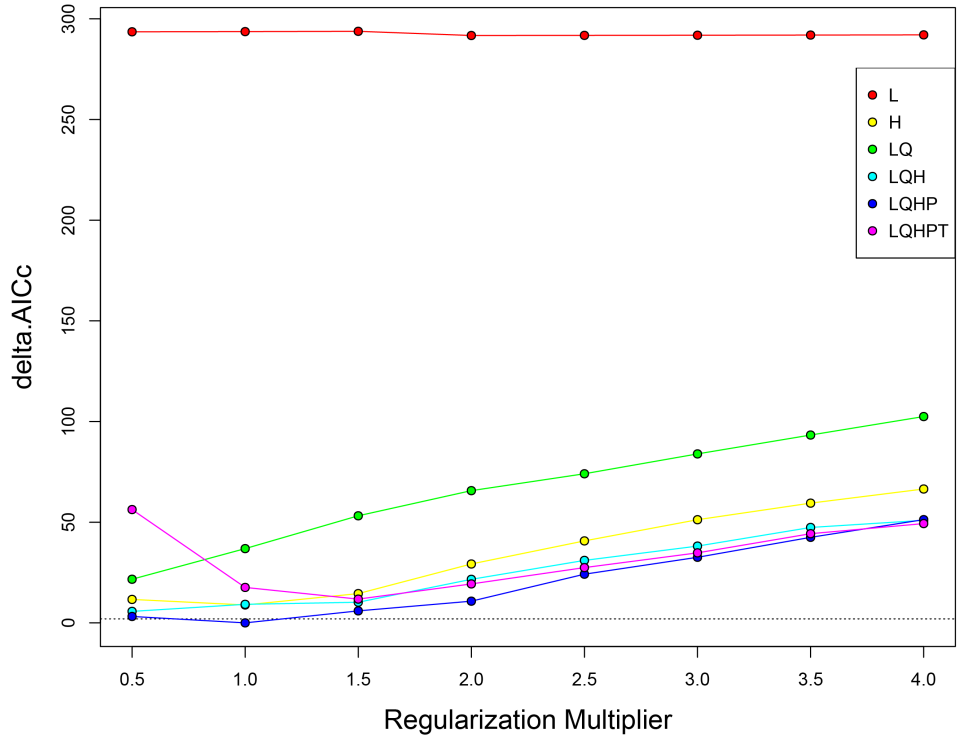

Figure S5. Delta AICc values of all models compared in ENMeval, model LQHP 1 was the best fit model.

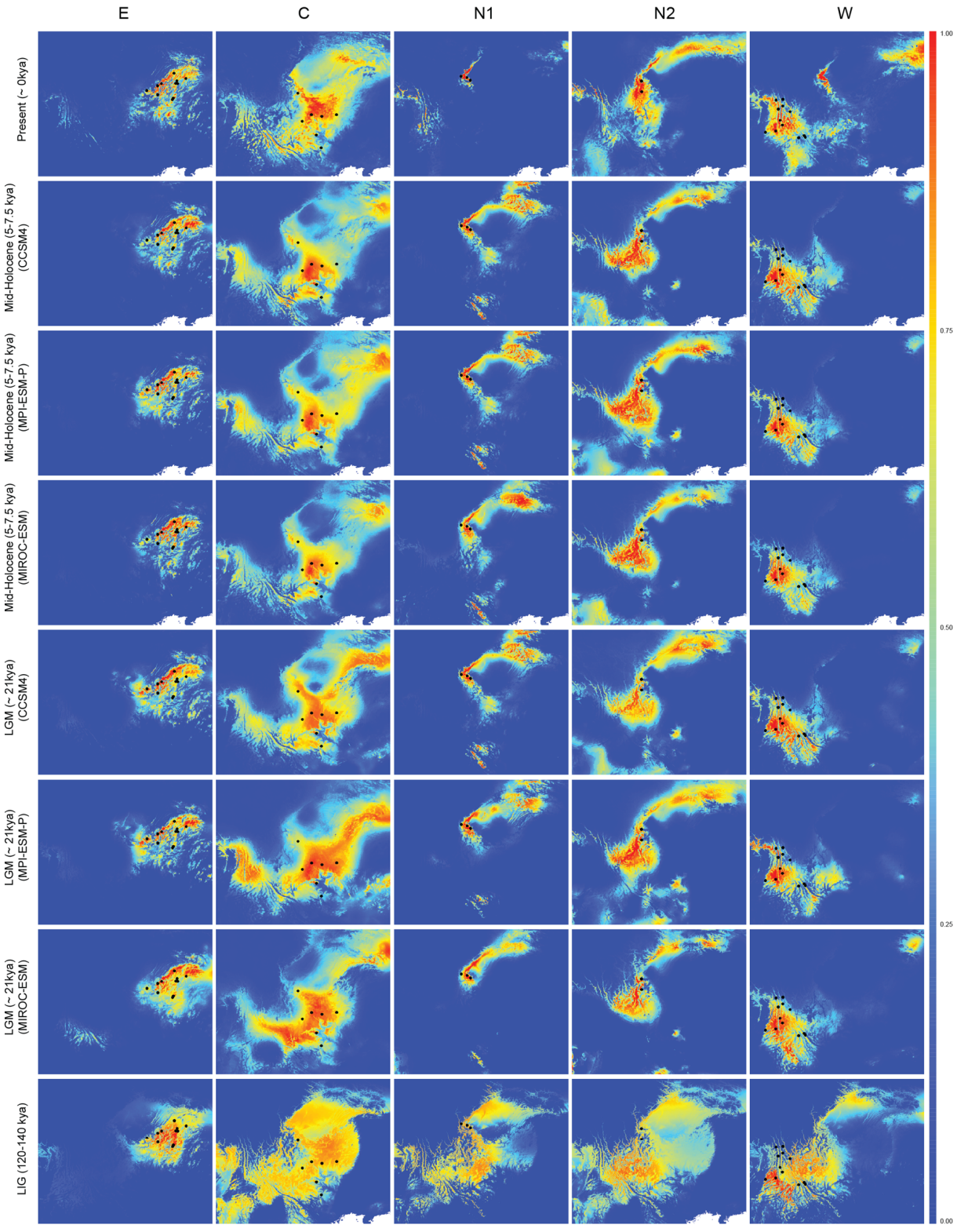

Figure S6. SDM for other models.

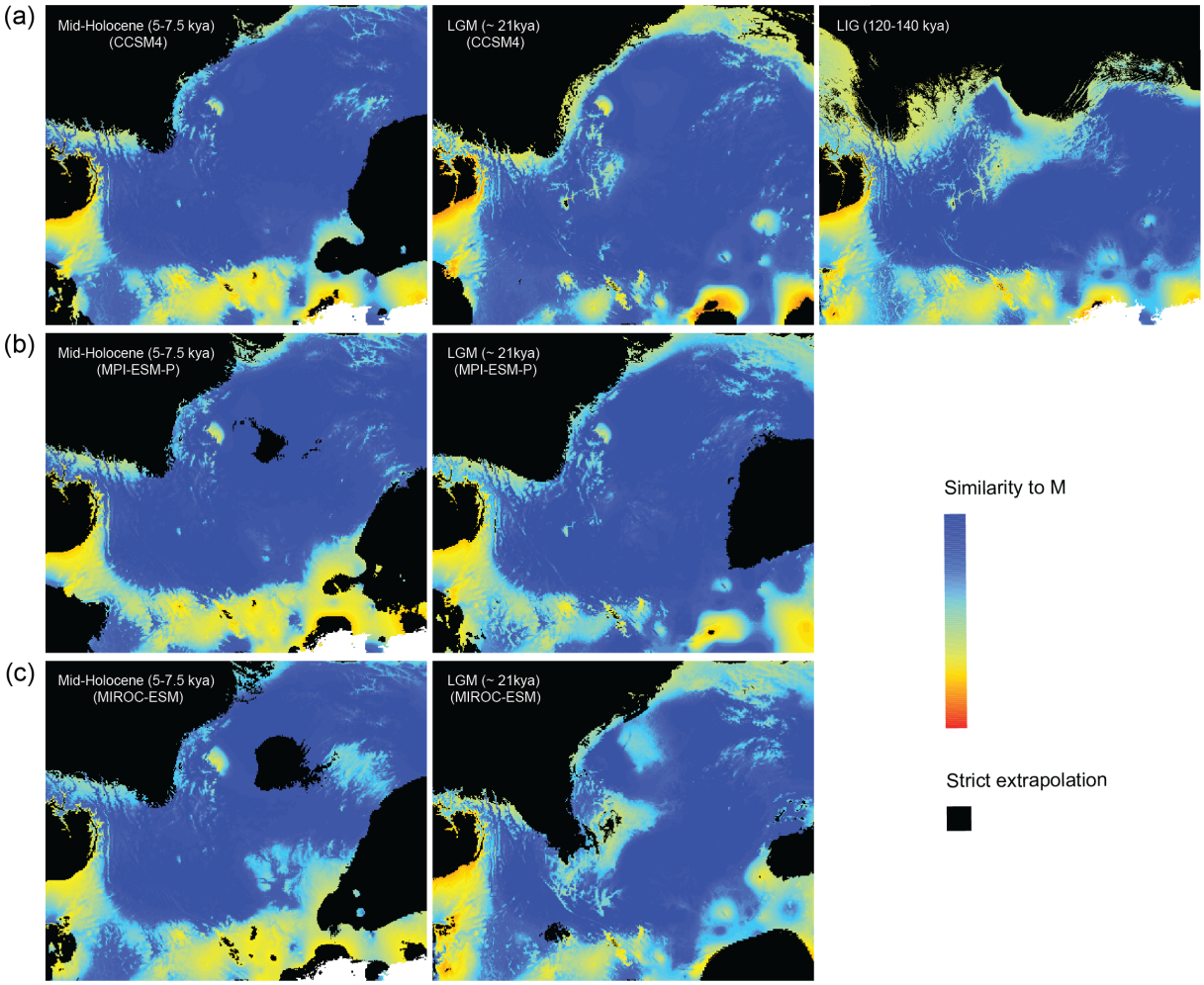

Figure S7. Mobility-oriented parity analysis comparing current conditions of the calibration region for *H. annectans* distribution modelling during three periods (MH, LGM, and LIG). (a) Results for the CCSM4 scenario. (b) Results for MPI-ESM-P and (c) the results for MIROC-ESM. Blue indicates similar climates to the current climate. Black indicates areas of strict extrapolation.

Table S9. Pairwise *F*-values of morphological features (snout-vent length [SVL] and weight) of 339 individuals from the seven g
were analyzed with Welch ANOVA using a Games-Howell *post hoc* test.

	SVL	Weight	SVL	Weight	SVL	Weight	SVL	Weight	SVL
Cluster C1	0,903	0,354							
Cluster C2	1,144	0.447*	0,240	0,093					
Cluster N1	0,160	0,333	1,064	0.687*	1,304	0.78*			
Cluster N2	0,215	0,263	1,119	0,617	1,359	0.71*	0,055	0,070	
Cluster W1	0,629	0,222	1,533	0.576*	1.773*	0.669*	0,469	0,111	0,415
Cluster W2	0,052	0,020	0,956	0,373	1,196	0.467*	0,108	0,313	0,163

* $P < 0.05$

enetic lineages of *H. annectans*. Data

Weight	SVL	Weight
0,040		
0,243	0,578	0,202